



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 163761

TO: Patricia Duffy
Location: 3b05 / 3c18
Sunday, August 28, 2005
Art Unit: 1645
Phone: 571-272-0855
Serial Number: 09 / 868195

From: Jan Delaval
Location: Biotech-Chem Library
Remsen 1a51
Phone: 571-272-2504

jan.delaval@uspto.gov

Search Notes

This Page Blank (uspto)

From: Duffy, Patricia
Sent: Thursday, August 25, 2005 8:33 AM
To: STIC-Biotech/ChemLib
Subject: sequence search 09/868195

In re: 09/868,195

Please search SEQ ID NO:12.

Please search commercial and interference databases.

Please print out top 75 hits in each category.

Thank you.

Patricia A. Duffy, Ph.D.
Art Unit 1645
Remsen 3B05; Mailbox 3C18
571-272-0855

STAFF USE ONLY

Searcher: Jan
Searcher Phone: 2- 2504
Date Searcher Picked up: 8/28/05
Date Completed: 8/28/05
Searcher Prep/Rev. Time: 0
Online Time: 12

Type of Search

NA#: _____ AA#: ✓
Interference: ✓ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: ✓
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 28, 2005, 10:18:29 ; Search time 41 Seconds

(without alignments)
1114.706 Million cell updates/sec

Title: US-09-868-195-12

Perfect score: 2393
Sequence: 1 LRKEYQNYNGEMKSSVNOI.....GIRYSIEATNTKSYLDWK 475

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 80 summaries

Database :

1: PIR 79:*
2: pir1:*
3: pir2:*
4: pir3:*
5: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1812	75.7	474	2	E95129	hypothetical protee
2	1802	75.3	474	2	D98000	glyceraldenhyde-3-p
3	1702	71.1	475	2	A57151	glyceraldenhyde-3-p
4	1383	57.8	482	2	H97348	NADP-dependent gly
5	1244	52.0	481	2	E83929	NADP-dependent gly
6	1168.5	48.8	496	2	F84634	hypothetical prote
7	1164.5	48.7	498	1	S43833	glyceraldenhyde-3-p
8	1143.5	47.8	496	1	S43832	glyceraldenhyde-3-p
9	868.5	36.3	475	2	F82900	NADP-dependent gly
10	763	31.9	488	2	G69752	NADP-dependent gly
11	753.5	31.5	463	2	B64476	NADP-dependent gly
12	752	31.4	470	2	F90347	hypothetical prote
13	750	31.3	470	2	D90323	hypothetical prote
14	741.5	31.0	455	2	H69230	NADP-dependent gly
15	741.5	31.0	489	2	A83033	probable aldenhyde
16	737	30.8	476	2	A70318	aldehyde denhydroge
17	736.5	30.8	491	2	T46684	p-hydroxybenzalden
18	735	30.7	482	2	B83867	NADP-dependent gly
19	726	30.3	482	2	F65045	succinate-semialde
20	721.5	30.2	488	1	I39769	aldehyde denhydroge
21	720	30.1	488	2	A11546	succinate semialde
22	718	30.0	488	2	A11188	succinate semialde
23	717	30.0	482	2	D85913	succinate-semialde
24	717	30.0	482	2	B91669	succinate-semialde
25	713.5	29.8	480	2	C84250	glyceraldenhyde-3-p
26	713	29.8	489	2	T44939	glyceraldenhyde-3-p
27	710.5	29.7	489	2	C84401	aldehyde denhydroge
28	710.5	29.7	498	2	AD2277	aldehyde denhydroge
29	697.5	29.1	468	2	C83774	succinate-semialde

ALIGNMENTS

30	682.5	28.5	485	2	B83901	aldehyde dehydroge
31	682	28.5	482	2	A80839	succinate-semialde
32	682	28.5	496	2	T44987	aldehyde dehydroge
33	673	28.1	485	2	C69584	aldehyde dehydroge
34	669.5	28.0	470	2	A87547	vanillin dehydroge
35	668	27.9	480	2	D83613	succinate-semialde
36	666.5	27.9	524	2	E75574	aldehyde dehydroge
37	666	27.8	509	1	S14629	aldehyde dehydroge
38	664.5	27.8	495	2	C89778	hypothetical prote
39	664	27.7	477	2	A83469	aldehyde dehydroge
40	661.5	27.6	509	2	E90504	hypothetical prote
41	661.5	27.6	509	2	E96825	hypothetical prote
42	661	27.6	517	1	A40872	aldehyde dehydroge
43	660	27.6	483	2	E49343	salicylaldehyde de
44	658.5	27.5	490	1	S15181	betaine-aldehyde d
45	658	27.5	501	1	DEH081	aldehyde dehydroge
46	656	27.4	485	2	C95964	probable aldehyde
47	655	27.4	487	2	AD3300	succinate-semialde
48	653.5	27.3	490	2	B85524	NADP-dependent bet
49	653	27.3	477	2	A95990	probable dehydroge
50	651.5	27.2	485	2	H72562	probable aldehyde
51	651.5	27.2	487	2	G75592	succinate-semialde
52	650	27.2	490	2	E95316	succinate-semialde
53	650	27.2	512	1	A55684	aldehyde dehydroge
54	649.5	27.1	490	2	AD0143	aldehyde dehydroge
55	649	27.1	496	2	T50272	betaine aldehyde d
56	649	27.1	996	2	AC1874	probable aldehyde
57	648.5	27.1	490	2	F90673	1-pyrroline-5 carb
58	647.5	27.1	495	2	A75608	NADP-dependent bet
59	647	27.0	495	1	G64878	aldehyde dehydroge
60	644.5	26.9	462	2	D69764	aldehyde dehydroge
61	643.5	26.9	485	2	C69814	succinate-semialde
62	643.5	26.9	490	2	C83675	benzaldehyde dehyd
63	643	26.9	495	2	E90863	glycine betaine al
64	643	26.9	495	2	D85755	aldehyde dehydroge
65	640	26.7	462	2	T36807	aldehyde dehydroge
66	639	26.7	484	2	A83176	probable aldehyde
67	639	26.7	501	2	S14752	NAD-dependent succ
68	639	26.7	990	2	S75910	aldehyde dehydroge
69	637.5	26.6	485	2	E83775	probable proline d
70	637.5	26.6	486	2	T36346	aldehyde dehydroge
71	637	26.6	479	2	T31293	probable aldehyde
72	636	26.6	484	2	F98208	aldehyde dehydroge
73	636	26.6	484	2	AC3078	succinate-semialde
74	635	26.5	499	2	S74224	aldehyde dehydroge
75	632.5	26.4	483	2	H98159	aldehyde dehydroge
76	632.5	26.4	483	2	A13127	hypothetical prote
77	630.5	26.3	497	1	A29055	vanillin: NAD oxid
78	630	26.3	497	2	B82981	aldehyde dehydroge
79	628.5	26.3	480	2	A81023	probable aldehyde
80	627.5	26.2	480	2	F81965	aldehyde dehydroge
						lactaldehyde dehyd

RESULT 1

hypothetical protein SP119 [imported] - Streptococcus pneumoniae (strain TIGR4)

C/Spectrum: Streptococcus pneumoniae

C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Jul-2004

C/Accession: E95129

R/Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Umayam, L.A.; White, O.; Salberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,

neon, T.; Hickey, B.K.; Holt, I.E.

Science 293, 498-506, 2001

A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A/Reference number: A95000; MUID:21357209; PMID:11463916

A/Accession: E95129

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-474 <KUR>

A:Cross-references: UNIPROT:Q97Q70; GB:AE005672; PIDN:AAK75230.1; PID:g14972596; GSPDB:G
A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP119

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 75.7%; Score 1812; DB 2; Length 474;

Best Local Similarity 74.1%; Pred. No. 7.6e-108;

Matches 349; Conservative 57; Mismatches 65; Indels 0; Gaps 0;

QY 5 YQNVNGEKSSVNOQIELSPIDSSLGVPAMTREVVDHAKAGREALPAMALTVER 64

DB 4 YQNLVNGKWKSSQEOITIVSPINOELGTVPAMTQTEADEAMQAPAAALPARALSAVER 63

QY 65 AQLYHKADIIERDKKEIATVLAKEISKAVNASVTEVVRTADLIRYAAEGIRLSTSADE 124

DB 64 AAYLHKTAALIERDKKEIGITLLAKEVAKGIKALIGEVRTADLIRYAAEGIRLSTQAAHE 123

QY 125 GGRMDASTGKLAIVIRQPVGIVLAIAPYNPVNLGSKIAIPALIGNVNMFKEPPTQGSV 184

DB 124 GGGFEATSKNKLAVRREPVGIVLAIAPFNYPVNLASAKIAPALIGNVNMFKEPPTQSSI 183

QY 185 SGLVLAKEAFARAGLPAGVFNITTRGSGEIGDIYVEHEVNFINTGSPVQORIGKLAGM 244

DB 184 SGLLAKAFEEBAGIPAGVFNITTRGSGEIGDIYIEHKEVNFINTGSPVQORIGKLAGM 243

QY 245 RPIMLEGKDGAGIVLADADLDNAKQIVAGAYDVSQGRCTAIRVLYVEEVADELAEKI 304

DB 244 RPIMLEGKDALVLEADLEHAAQIVAGAFYSISQGRCTAIRVLYVESVADKATLTL 303

QY 305 SENVAKLSVGDPPDNATVPIIDNSADPIESLVVDARQKAKELNEFKRDGRLLTPGLF 364

DB 304 QEVSRLTVGDPFDNADITPVIDNASADPIWGLIEDAQEKBAQALPIKREGNLLMPVL 363

QY 365 DHVTLDMKLAMEBPPGPIPIIRYKDAEAAVAIANKSDGLQSSVTRPFOKAFDIANKL 424

DB 364 DQVTKDMKVAMEBPPGPVLPPIIRVASVEBAIAFANSEBFGLOSSVFTNDFKKAFAIEAKL 423

QY 425 EVGTVINNKTRGPDNPFFLGLKSGAGVQIGIRYSIEMTNVKSIVLDMK 475

DB 424 EVGTVINNKTRGPDNPFFLGVKSGAGVQIGIKYSIEMTNVKSIVPDK 474

RESULT 2

D98000

glyceraldehyde-3-phosphate dehydrogenase (NADP) (EC 1.2.1.9) gapN [imported] - Streptococ

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C:Accession: D98000

R:Host(s): J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; F

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.;

A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: D98000

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-474 <KUR>

A:Cross-references: UNIPROT:Q8DP87; GB:AE007317; PIDN:AAK99832.1; PID:g15458646; GSPDB:G

C:Genetics:

A:Gene: gapN

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

C:Keywords: oxidoreductase

Query Match 75.3%; Score 1802; DB 2; Length 474;

Best Local Similarity 73.7%; Pred. No. 3.3e-107;

Matches 347; Conservative 58; Mismatches 66; Indels 0; Gaps 0;

QY 5 YQNVNGEKSSVNOQIELSPIDSSLGVPAMTREVVDHAKAGREALPAMALTVER 64

DB 4 YQNLVNGKWKSSQEOITIVSPINOELGTVPAMTQTEADEAMQAPAAALPARALSAIER 63

QY 65 AQLYHKADIIERDKKEIATVLAKEISKAVNASVTEVVRTADLIRYAAEGIRLSTSADE 124

DB 64 AAYLHKTAALIERDKKEIGITLLAKEVAKGIKALIGEVRTADLIRYAAEGIRLSTQAAHE 123

QY 125 GGRMDASTGKLAIVIRQPVGIVLAIAPYNPVNLGSKIAIPALIGNVNMFKEPPTQGSV 184

DB 124 GGGFEATSKNKLAVRREPVGIVLAIAPFNYPVNLASAKIAPALIGNVNMFKEPPTQSSI 183

QY 185 SGLVLAKEAFARAGLPAGVFNITTRGSGEIGDIYVEHEVNFINTGSPVQORIGKLAGM 244

DB 184 SGLLAKAFEEBAGIPAGVFNITTRGSGEIGDIYIEHKEVNFINTGSPVQORIGKLAGM 243

QY 245 RPIMLEGKDGAGIVLADADLDNAKQIVAGAYDVSQGRCTAIRVLYVEEVADELAEKI 304

DB 244 RPIMLEGKDALVLEADLEHAAQIVAGAFYSISQGRCTAIRVLYVESVADKATLTL 303

QY 305 SENVAKLSVGDPPDNATVPIIDNSADPIESLVVDARQKAKELNEFKRDGRLLTPGLF 364

DB 304 QEVSRLTVGDPFDNADITPVIDNASADPIWGLIEDAQEKBAQALPIKREGNLLMPVL 363

QY 365 DHVTLDMKLAMEBPPGPIPIIRYKDAEAAVAIANKSDGLQSSVTRPFOKAFDIANKL 424

DB 364 DQVTKDMKVAMEBPPGPVLPPIIRVASVEBAIAFANSEBFGLOSSVFTNDFKKAFAIEAKL 423

QY 425 EVGTVINNKTRGPDNPFFLGLKSGAGVQIGIRYSIEMTNVKSIVLDMK 475

DB 424 EVGTVINNKTRGPDNPFFLGVKSGAGVQIGIKYSIEMTNVKSIVPDK 474

RESULT 3

A57151

glyceraldehyde-3-phosphate dehydrogenase (NADP) (EC 1.2.1.9) - Streptococcus mutans

C:Species: Streptococcus mutans

C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 09-Jul-2004

C:Accession: A57151

R:Boyd, D.A.; Cviltkovitch, D.G.; Hamilton, I.R.

J. Bacteriol. 177, 2622-2727, 1995

A:Title: Sequence, expression, and function of the gene for the nonphosphorylating, NADP

A:Reference number: A57151; MUID:95270576; PMID:7751269

A:Accession: A57151

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-475 <BOY>

A:Cross-references: UNIPROT:Q59931; GB:IJ38521; GB:IJ5191; NID:g642666; PIDN:AAA91091.1;

C:Genetics:

A:Gene: gapN

A:Start codon: TTG

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

C:Keywords: NADP; oxidoreductase

F:41-304/Domain: aldehyde dehydrogenase homology <AND>

Query Match 71.1%; Score 1702; DB 2; Length 475;

Best Local Similarity 66.9%; Pred. No. 7.6e-101;

Matches 318; Conservative 75; Mismatches 82; Indels 0; Gaps 0;

QY 1 LTKYQNVNGEKSSVNOQIELSPIDSSLGVPAMTREVVDHAKAGREALPAMALT 60

DB 1 MTKQYKXNVNGEKSSVNOQIELSPIDSSLGVPAMTREVVDHAKAGREALPAMALT 60

QY 61 YVERAQYHLKADIIERDKKEIATVLAKEISKAVNASVTEVVRTADLIRYAAEGIRLST 120

DB 61 YIERAAIYHLKADIIERDKKEIATVLAKEISKAVNASVTEVVRTADLIRYAAEGIRLST 120

QY 121 SADGGKMDASTGKLAIVIRQPVGIVLAIAPYNPVNLGSKIAIPALIGNVNMFKEPPT 180

DB 121 EYLGGSEFAASKKIIVRREPVGIVLAIAPFNYPVNLASAKIAPALIGNVNMFKEPPT 180

QY 181 QGSVGLVLAKEAFARAGLPAGVFNITTRGSGEIGDIYVEHEVNFINTGSPVQORIGK 240

DB 181 QGISGLLAEAFARAGLPAGVFNITTRGSGEIGDIYVEHEVNFINTGSGEIGERIGK 240

QY 241 LAGMRPIMLELGKDGAGIVLADADLDNAKQIVAGAYDVSQGRCTAIRVLYVEEVADEL 300

A:Residues: 1-496 <STO>
A:Cross-references: UNIPROT:Q9ZUG8; GB:AE002093; NID:g4115387; PIDN:AAD03388.1; GSPDB:GN
C:Genetics:
A:Gene: At2g24270
A:Map position: 2
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 48.8%; Score 1168.5; DB 2; Length 496;
Best Local Similarity 50.3%; Pred. No. 7,2e-67;
Matches 240; Conservative 84; Mismatches 142; Indels 11; Gaps 5;

QY 5 YQNYNGEKK--SSVNOEILSPIDSSLGFPVPMATREVDHAKAGREALPAMALTY 62
DB 16 YKYVADGEMKSSSGSVAINMPATRKTYQVACQOEENAVMEIAKSAQCKMAKTPLM 75
QY 63 BRAQYLHKAADIIEEDKEIATVLAKEISKAVNASVTEVTRADLIRYAEGRILSTSA 122
DB 76 KRAELHKAALIKENKKAIAECLVKEIAKPAKDAVSEVRSGLVSYTAEGVRILGE- 132
QY 123 DEGGKM--DASTGH--KLAVIRROPVGIYLAIAPNYPVNLSSGIAPALIGVNMFK 177
DB 133 GEGKFLVSDSPFGNERTKCYCLTSKIPLGIVLAIPEPNYPVNLAVSKIAPALIAGNSIVLK 192
QY 178 PPTQSGVGLVLAKAFPAEAGLPAGVPNTITGRGSEIGDIYVEHEVNFINTGSTPVQOR 237
DB 193 PPTQAVSCILHNVHCFHLAEPKGLISCTTGKSGEIGDPLTMHPAVNCISFTGG-DTGIS 251
QY 238 IGLAGMPIMLELGGKAGIYVLADADLNAKQIVAGAYDYSGORCTAIRVLVEEVA 297
DB 252 ISKAGKMTPLQMELGKDACIVLDHDLVANSNIIKGSFYSQGCCTAVKVLVMESEA 311
QY 298 DELAEKISENVAKLSVGDPDFNATVTPVIDNSADFIESLVVDAROKAKELNEFRKDR 357
DB 312 DELVEKIVAKVAKLTGVPPEENSDITAVVSESSANFIEGLVMDAKKATPCQETKREBN 371
QY 368 LITPGLFDHYTLDMKLAMEEPFGPIIPIIRVKDAEAVAIANKSDPGLQSSVPTRDFOK 417
DB 372 LITPGLFDHYTLDMKLAMEEPFGPIIPIIRVKDAEAVAIANKSDPGLQSSVPTRDFOK 431
QY 418 FDIANKLEVGTVHINNKTRGPDNFPPLGLKSGAGVQIRYSIEAMTNVKSIVLDM 474
DB 432 ILISDAMETGTVOINSAPARGPDPHFSPQGLKDSIGSGQVITNSINMTTKVKSIVINL 488

RESULT 7

543833
glyceralddehyde-3-phosphate dehydrogenase (NADP) (EC 1.2.1.9) - maize
C:Species: Zea mays (maize)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S43833
R:Habenicht, A.; Hellman, U.; Cerff, R.
J. Mol. Biol. 237, 165-171, 1994
A:Title: Non-phosphorylating GAPDH of higher plants is a member of the aldehyde dehydrog
A:Reference number: S43832; MUID:94180387; PMID:7545914
A:Accession: S43833
A:Molecule type: mRNA
A:Residues: 1-498 <HAB>
A:Cross-references: UNIPROT:Q43272; EMBL:X75326; NID:g474407; PIDN:CAA53075.1; PID:g4744
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: NADP; oxidoreductase
F:56-320/Domain: aldehyde dehydrogenase homology <ALDD>
F:266,300/Active site: Glu, Cys #status predicted

Query Match 48.7%; Score 1164.5; DB 1; Length 498;
Best Local Similarity 49.0%; Pred. No. 1.3e-66;
Matches 234; Conservative 88; Mismatches 143; Indels 13; Gaps 5;

QY 5 YQNYNGEKKSVN-QIEILSPIDSSLGFPVPMATREVDHAKAGREALPAMALTY 62
DB 18 YKYVADGEMKSSSGSVAINMPATRKTYQVACQOEENAVMEIAKSAQCKMAKTPLM 77
QY 63 BRAQYLHKAADIIEEDKEIATVLAKEISKAVNASVTEVTRADLIRYAEGRILSTSA 122

DB 78 KRAVTLHKAALIKENKKAIAECLVKEIAKPAKDAVSEVRSGLVSYTAEGVRILGE- 136
QY 123 DEGGKM--DASTG--HKLAVIRROPVGIYLAIAPNYPVNLSSGIAPALIGVNMFK 176
DB 137 --GKLVSDSPFGNERNKCYCLTSKIPLGIVLAIPEPNYPVNLAVSKIAPALIAGNSIVLK 193
QY 177 KPTQSGVGLVLAKAFPAEAGLPAGVPNTITGRGSEIGDIYVEHEVNFINTGSTPVQO 236
DB 194 KPTQGAVALHNVHCFHLAEPKGLISCTTGKSGEIGDPLTMHPAVNCISFTGG-DTGI 252
QY 237 RIGKLAKRPMLELGGKAGIYVLADADLNAKQIVAGAYDYSGORCTAIRVLVEEVA 296
DB 253 AISKAGVMPLOMELGKDACIVLEDADLVSANITVGGFSYSGORCTAVKVLVIMESI 312
QY 297 DELAEKISENVAKLSVGDPDFNATVTPVIDNSADFIESLVVDAROKAKELNEFRKDR 356
DB 313 ADAVOKVNAKALIKVGPDDSDITPVVSESSANFIEGLVMDAKKATPCQETKREBG 372
QY 357 RUTPGLFDHYTLDMKLAMEEPFGPIIPIIRVKDAEAVAIANKSDPGLQSSVPTRDFOK 416
DB 373 NLWPLLDHYRPMRIAMEEPFGVLPVIRINSVEGIHHCNANFGLQGCIFTRDINK 432
QY 417 AFDIANKLEVGTVHINNKTRGPDNFPPLGLKSGAGVQIRYSIEAMTNVKSIVLDM 474
DB 433 AILSDAMETGTVOINSAPARGPDPHFSPQGLKDSIGSGQVITNSINMTTKVKSIVINL 490

RESULT 8

543832
glyceralddehyde-3-phosphate dehydrogenase (NADP) (EC 1.2.1.9) - garden pea
C:Species: Pisum sativum (garden pea)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 15-Mar-2004
C:Accession: S43832
R:Habenicht, A.; Hellman, U.; Cerff, R.
J. Mol. Biol. 237, 165-171, 1994
A:Title: Non-phosphorylating GAPDH of higher plants is a member of the aldehyde dehydrog
A:Reference number: S43832; MUID:94180387; PMID:7545914
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-496 <HAB>
A:Cross-references: EMBL:X75327
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: NADP; oxidoreductase
F:54-318/Domain: aldehyde dehydrogenase homology <ALDD>
F:264,298/Active site: Glu, Cys #status predicted

Query Match 47.8%; Score 1143.5; DB 1; Length 496;
Best Local Similarity 48.8%; Pred. No. 2.8e-65;
Matches 233; Conservative 85; Mismatches 148; Indels 11; Gaps 5;

QY 5 YQNYNGEKKSVN-QIEILSPIDSSLGFPVPMATREVDHAKAGREALPAMALTY 62
DB 16 YKYVADGEMKSSSGSVAINMPATRKTYQVACQOEENAVMEIAKSAQCKMAKTPLM 75
QY 63 BRAQYLHKAADIIEEDKEIATVLAKEISKAVNASVTEVTRADLIRYAEGRILSTSA 122
DB 76 KRAELHKAALIKENKKAIAECLVKEIAKPAKDAVSEVRSGLVSYTAEGVRILGE- 132
QY 123 DEGGKM--DASTGH--KLAVIRROPVGIYLAIAPNYPVNLSSGIAPALIGVNMFK 177
DB 133 GEGKFLVSDSPFGNERTKCYCLTSKIPLGIVLAIPEPNYPVNLAVSKIAPALIAGNSIVLK 192
QY 178 PPTQSGVGLVLAKAFPAEAGLPAGVPNTITGRGSEIGDIYVEHEVNFINTGSTPVQOR 237
DB 193 PPTQAVVALHNVHCFHLAEPKGLISCTTGKSGEIGDPLTMHPAVNCISFTGG-DTGI 251
QY 238 IGLAGMPIMLELGGKAGIYVLADADLNAKQIVAGAYDYSGORCTAIRVLVEEVA 297
DB 252 ISKSGKMTPLQMELGKDACIVLEBDADLVSANIIKGSFYSQGCCTAVKVLVMESEA 311
QY 298 DELAEKISENVAKLSVGDPDFNATVTPVIDNSADFIESLVVDAROKAKELNEFRKDR 357

Db 312 DALVEKVKVAKLSVGPEDSDITPVSVSESSANFLBGLVNDAXEKGAFTQCEYKREBN 371
QY 358 LITPGFEDVHTLDMKLAWEPPFPIPIIRVKAEEAVAIANKSDGLQSSVTRPQKA 417
Db 372 LITPLLDVNRPRMIRAMEBPPFVPIVRINSVEEGIHHCNANSNGLQGCVCFTKIDINRA 431
QY 418 FDTANKLEVTGVINNKTKGRGNPPFLGLKSGAGVQIGRYSIEAMTWKSVILVM 474
Db 432 IMISDMESGTVINAPARGPDPHFSGQIKDSGISQGITNSINMTKVKITVINL 488

RESULT 9
F82900
NADP-dependent glyceraldehyde-3-phosphate dehydrogenase U0362 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Mar-2004
C:Accession: F82900
R:Glaes, J.I.; Leifkowitz, E.J.; Glaes, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mir
A:Reference number: A82870
A:Accession: F82900
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-475 <GLA>
A:Cross-references: GB:AE002133; GB:AF222894; NID:g66939339; PIDN:AAF30771.1; GSPDB:GN001
A:Experimental source: serovar 3; biovar 1
A:Genetics:
A:Gene: gapN; U0362
A:Genetic code: SGC3
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 36.3%; Score 868.5; DB 2; Length 475;
Best Local Similarity 39.7%; Pred. No. 8.4e-48;
Matches 190; Conservative 95; Mismatches 177; Indels 17; Gaps 5;

QY 5 YQNVYNGEKSSVNOQIEILSPIDSSLGFPVPMATRE-EVDHAKAGREALPAMALTYVE 63
Db 3 YKTLNGAFVDAEKKLPVNPSPNQIIVAVPNHNENEINTTFENAHIVLKEKDPPIKY 62
QY 64 RAQYIKAAIDIERDKEEIAITVLAKEISKAYNASVTEVTRADLIRYAAEGIRLSTSA- 122
Db 63 RCDLLKLADKDEHKOGLAQIISTEIAKGLKDSLEVERSAHYLAETVPEYQKLMQKRI 122
QY 123 --DEGKMDASTGH---KLAIVTRQPGVIGVIAIADYNTPVNLSGSKIAPALIGNVNMF 176
Db 123 IPDE-----TVHAYKNKVAATFYRIPVGVVLAICFPNYPINLLISKLPALVSGNSLVY 175
QY 177 KPPTQGSVSGVILAKAFAPAGLPAGVFNITTGSGSEIGDYVHEBVNINFTGSTPVGQ 236
Db 176 KSTQGSLLGIRISLVHVGFPKGVNCLTEARITGDLVTKVKVKAISFTGGSPKVN 235
QY 237 RIQGLKAMPIMLELGKAGIVLADADLDNNAKOIVAGAYDSGQRTAIRVLAEEV 296
Db 236 HIAETISKISLVHLEGGKOPALVLDADBELANETVKAIGYSGQRCTAIRVFPVSHQK 295
QY 297 ADELAKEISENVAKLSVGDPPDNATVPVVIDNSADFIESLVVDARQKA--KEINFKR 354
Db 296 HDLVLVNLINKVDALTVGLPQGNPITPLINSNLKYNLSLVEDAIKKGALVHQKLVNE 355
QY 355 DGRLLTPGLFDVHTLDMKLAWEPPFPIPIIRVKAEEAVAIANKSDGLQSSVTRPQ 414
Db 356 KNNLHPVLDVNTKMRVAMEBPPFPIPIITVYNSIQEADIDINSQYGLQACIPTTVY 415
QY 415 OKAFDIANKLEVTGVVHINKTKGRGNPPFLGLKSGAGVQIGRYSIEAMTWKSVILV 473
Db 416 ASTEQALQIESITININKSSSRGPDLPPFGVKGDFVQGLVDAIISMTTIKGIIN 474

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: G69752
R:Kunze, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Enllich, S.D.; Emerson, P.T.; Enlian, K.D.; Erington, J.; Fabret, C.; Ferrati, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallie
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulo, M.F
Koetler, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seru
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Teperstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbato, H.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:3384377
A:Accession: G69752
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-488 <KUN>
A:Cross-references: UNIPROT:P42236; GB:Z99105; GB:AL009126; NID:g2632457; PIDN:CAB12041
A:Experimental source: strain 168
C:Genetics:
A:Gene: ycbD
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
F:49-309/Domain: aldehyde dehydrogenase homology <ALD>

Query Match 31.9%; Score 763; DB 2; Length 488;
Best Local Similarity 36.9%; Pred. No. 4.5e-41;
Matches 170; Conservative 94; Mismatches 179; Indels 18; Gaps 6;

QY 5 YQNVYNGEW--KSSVNOQIEILSPID-SSLGFPVPMATRE-EVDHAKAGREALPAMALTYV 61
Db 10 YNPFNGEAVVKSQSDGMKVENPADVNDIVGVQNSTADVERAYTAANEATARKLTLG 69
QY 62 YRRAQYIKAAIDIERDKEEIAITVLAKEISKAYNASVTEVTRADLIRYAAEGIRLST 121
Db 70 AERGQVLYKTADIMEQRLEIAACATREWGKTLPEAKGETARQIAILRYAAGEGRMKT-- 127
QY 122 ADEGKMDASTGHKLAIVTRQPGVIGVIAIADYNTPVNLSGSKIAPALIGNVNMF 181
Db 128 ---GDVISTDQDALMFTTRVPLVSAVGVISPNFPALPYMKAAPALVAGNVVILKIPATE 184
QY 182 GSVSGVILAKAFAPAGLPAGVFNITTGSGSEIGDYVHEBVNINFTGSTPVGQRIKT 241
Db 185 TAVTCAKTIACFEBAGLPAGVINLVTPGSGVVGGLAHDGVNAVTFPGSNQVGIQGA 244
QY 242 AQMR--PIMLELGKAGIVLADADLDNNAKOIVAGAYDSGQRTAIRVLAEEV 299
Db 245 ALARGAKYOLEWGGKNPVIVADADLEAABAVIIGAFSTQCKTATSRTVIOGSIYER 304
QY 300 LAEKISENVAKLSVGDPP-DNATVPVVIDNSADFIESLVVDARQKAEL-----NE 351
Db 305 FKEKLTQTKDITDLSLEKEDVWMPGPIASKNDLDCLESTIEKQKQSGSLIGGEKLENG 364
QY 352 FRDGRLLTPGLFDVHTLDMKLAWEPPFPIPIIRVKAEEAVAIANKSDGLQSSVTRP 411
Db 365 KYQNGVYVQPAFDVNTSEMTIAQIEIFQPVIALIKVDISIEALMIANDVKRGLASITP 424
QY 412 RQFQAFDIANKLEVTGVVHINKTKGRGNPPFLGLKSGAG 452
Db 425 ENIGMLSFIDEIDAGLVINAESAGVELQAPFGMKOSSS 465

RESULT 11
B64476
NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.-) - *Methanococcus jan*
C:Species: *Methanococcus jannaschii*
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: B64476
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

A: Reich, C.I.; Overbeek, R.; Kitzman, E.F.; Weinstein, K.G.; Merrick, J.M.; Glodok, A.
Ison, U.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurel, M.A.
Citation: Science 273, 1058-1073, 1996
A: Authors: Kalne, B.P.; Borodovsky, M.; Kleen, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
C.R.; Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A: Accession number: Af43300; PMID:96337999; PMID:868087
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-463 <BDU>
C: Cross-references: UNIPROT:Q58806; GB:U67581; GB:L77117; NID:g1592052; PID:g1592060; T
K: Keywords: MADP; oxidoreductase
S: Superfamily: MAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
E: Query Match 31.5%; Score 753.5; DB 2; Length 463;
Best Local Similarity 36.9%; Pred. No. 1.7e+40;
Matches 174; Conservative 95; Mismatches 184; Indels 19; Gaps 8;

OY YNGEWSKSVNQ--IEILSPIDDSSLGFPVPMTRREVDHANKAGREALPAWALTYYERA 65
DB ::::|::|::::|:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
2 FIDGKM---INREDMDVINPYSLSEVIKKIPALSRBEAKAIDLTAKEYEKWMKNLPITRY 58
OY QYLKKAADIERDDEELTYVLAKEISKAYNSVTVEVPTADLIRAAEEGRISTSADEG 125
DB ::||:|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
59 NILNNIAKOIKKEKEELAKILAIDAGKPICKARFEVERSICTFKLA--FYVKHRDEV 115
OY 126 GKMDASTGHKLAVLRPGGIIVLAAPNVYPNLSGSIALAPLAGANVMFPPTOGSVS 185
DB :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
116 IPSD----DRLLFRREVGVIGAITTPNPPLNTSANKIALAIATANGVIYNHPSSKAPLV 171
OY 186 GLVLAK---APEARGDALVGVFNTTTGGSGESIGDIVYEHEEVNFINTFGSTPVGORIQL 241
DB :||:|:~::~|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
172 CIELAKITENLKRVNYVELGYVNLTGAGEVVGDENVNEKNMISFTGSSKGVELTIKK 231
OY 242 AGMPRIMLELGGKAGIYTLADADDNAKQIYAAGYDVDSGORCRALKVLYVEVADEL 301
DB :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
232 AGPKTIALETGLGVNPIVTKDAILKANVALIKGPSFIYAGVCISVGMITLVDESIADEKI 291
OY 302 EKISENWAKLTSGDPDFDNAT-VTPVIDDNSADPFISLTVDDARKOKAEKLNFFKRDLTLT 360
DB EMFNVRKAVLVGNPNLDKETDGVPRLISVENHAWEKVEEKAIMDGGSKILLGGKDHALFY 351
OY 361 PGLFDPHTVLTDMKLMEEPFGPILPIIRVKDAEEAVAIVANSKDPFGLOSSVFTRDPKAFDI 420
DB 352 PTILE-YDRNMILTCTETRFAPVIPITRTNB-EEMIDIANSRGTEYGHSAMIFINDINKSLKF 409
OY 421 ANKLEVGVTHINNKTGRPNDNPFPIGLKSGAGACOGVIGYSIEBMTNVKSITYL 472
DB 410 AENNLEFGGVINDSSLFRQDNMFPGVKKSGGLGREGVKYAMEBSMNIXYTIIII 461

RESULT 12
P90347
hypothetical protein gapn-2 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: P90347
R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, V.; Allard, G.; Sawyer, M.J.; Chan,
Jong, I.; Jeffries, A.C.; Kozaeta, C.D.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,
aretz, R.A.; Ragani, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: P90347
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-470 <KUR>
A:Cross-references: UNIPROT:Q97XA5; GB:AEO06641; NID:g13815099; PID:NAAK42037.1; GS:PDB:
C:Genetics:
A:Gene: gapn-2
C:Superfamily: MAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 31.4%; Score 752; DB 2; Length 470;
Best Local Similarity 36.5%; Pred. No. 2,2e-40;
Matches 172; Conservative 96; Mismatches 187; Indels 16; Gaps 7;

OY SVAOIELISPIDDSLGFVPAMTREEDVHAMKRGREALPAMAALTVERAQLYHKAADIT 75
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB SLVEIVRSPBNLKVIQTVKRMSKDEVRGEIEEAYKGFELISSMPLYKRAILRKISEIL 62
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY ERDKEIATVLAKESIAFNAASVTEVVRTADLRYYAAE-GIRLISTADEGKKMDA--- 130
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB EREGERLARLLAMEAGKPIDSRSRYEVLRASRLRQAFAEEGVILE--GKNYRVDAVEYP 119
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY -STGHKLAVIRROQVGIVLAIPAIFYPNVLSSGKIAPALIGGVNMEKPPQTQSVSGLVL 189
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB PGNENRIYLSREPIGVVLTALIPNEFPINSFAHVAPALAVGNSVVKPSINTPLAIVEN 179
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY AKAFEAAGLPAGVENNTTTRGRSEIGDIYVHEEYNFINFGSTPVGORIGKLA---GMRP 246
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB KKLIVEAGLPDSANARVVTGYSSSEIGDEIIHPLVGLITLNGSQTGKIKSXAISLGR- 238
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY TMLSELGGKAGIVLAADDLNAAKOIYAGKYDYSGQRCTAIKXVLVVEEYADELAEKISTE 306
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB IIMELGSDPFIILEDANIERASSIAVARAFEYAGQNCNAGKRIYVOEYVDKRFVAFNE 298
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY NVALLSYGDPEDNAT-VTPYIDNKSADFIESLVDVANOKGAKE--INEFRRDRLLTPGL 363
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB KAKALKIGEPEIDETTDVGPINKESVENLANSYLEDAKVKGGRBEILNRGESSESPFLTM 358
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY FDHTVTLDMKLAMEBPFEPILPIIRVKDAEBAVALANKSDFGLOSVFTRDFOKAFDIANK 423
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB VTNPSLDMLYLKSEVFPGPIVPIVSKSDEFAIRIANSTEGLOSAIPTNDVNRLKLSRE 418
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY LEVGTTHHNKTGRGPDNPPEPLGLKSGGAGVQGIRSIEMNTVKSIVLDM 474
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB LKFQAVLIINDSTRLRWDSLPRFGFKKTIGREGVRETMLEMTENKLIATIL 469
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 13

D90323
[hypothetical protein gapN-1 [imported] - Sulfolobus solfataricus
C|Species: Sulfolobus solfataricus
C|Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C|Accession: D90323
R|Jong, I.; Jeffries, A.C.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.U.; Chan-
Jung, R.A.; Ragen, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A|Description: Sulfolobus solfataricus complete genome.
A|Reference number: A99139
A|Accession: D90323
A|Status: preliminary
A|Molecule type: DNA
A|Reidsues: 1-470 <KOR>
A|Cross-references: UNIPROT:Q97XS9; GB:AEO06641; NID:G13814871; PIDN:AAK41843.1; GSPDB:C-
C|Genetics:
A|Gene: gapN-1
C|Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 31.3%; Score 750; DB 2; Length 470;
Best Local Similarity 36.6%; Pred. No. 2.9e-40;
Matches 171; Conservative 94; Mismatches 186; Indels 16; Gaps 7;

OY IEILSPIDDSSLGFPAMTREEDVHAMKAGREALPMAALTVERAQLYHKAADIIRDK 79
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB IEVSPSLNVLTIGTKRMHODEVRGEIEEAYKGEIITSKMPLYKRAILRKVSIIBEREQ 66
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY EEIATVLAKESIAFYNASVTEVVRTADLRYYAAE-GIRLISTADEGKKMDA-----STG 133
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB ERLAKTLAMEWGKRIQRSRVLEVLRASRLRPHAAFEVGMVLE--GKNYRVDAVEYPPGNE 123
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY HCLAVIRPOVGIYLAIPAIFYPNVLSSGKIAPALIGGVNMEKPPQTQSVSGLVLAKAF 193
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Db 124 NRIVSTREPIGVVTAALPFNPINSPAHKAVDAPALVAGNSVVKPSINTPLAITEMKIL 183
 Qy 194 AEEGLPAGVNTTITGKSEIGDYIVHEEVPNFIPGSGTVGQRIQGLA--GMRPLME 250
 Db 184 VEEGLPDSAVRVVTVGSSEIGDEIITHPLVGLITLIGSTGTGKIKASKAVSLKR-IIME 242
 Qy 251 LGGKDAIGIVLADDLNAAKOIVAGAYDYGORCTAIKRVLYVEEVDDELAEKISENVAK 310
 Db 243 LGGSDPIIIEEDNANIEREASSIAVRARFETAGQNCNCKRIITVEEYDKVAKAFNDKARA 302
 Qy 311 LSVGDPEDNAT-VTPVIDNSADFIESTLVVDAROKAKE--LNEPKRGSLITPGLFDHV 367
 Db 303 LKVGDLDETTDVGPIVINKSEVENINSVLADAKAGKVEILNKGPSSGFFPLTWVTPN 362
 Qy 368 TLDMKAMEEPFGPIPLIRVYKDAEVAIAJANKSDFGLOSSVTRDFOKAFDIANKLEVQ 427
 Db 363 SLMLVLKSEFVPIPIVIVSVKDEEAIKIANSTEVGLQGAIFPDNDVRALKLSRELKFG 422
 Qy 428 TVHINKTGGRPNFPLGLKSGAGVQIGRIYSIEAMTNKSIYLDK 474
 Db 423 AVIINSTRLRWDSLAFGFGFKSSIGREGVETLWENTENKLIATIL 469

RESULT 14

H69230
 NADP-dependent glyceraldehyde-3-phosphate dehydrogenase - Methanobacterium thermoautotrophicum
 C/Species: Methanobacterium thermoautotrophicum
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 A/Accession: H69230
 R/Smith, D.R.; Doucette-Stamm, L.A.; Delouche, C.; Lee, H.; Dubois, J.; Alredge, T.;
 Otu, D.; Spadecora, R.; Vitacore, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
 Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Moelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A/Reference number: A69000; MUID:98037514; PMID:9371463
 A/Accession: H69230
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-455 <MTH>
 A/Cross-references: UNIPROT:Q27059; GB:AE000871; GB:AE000666; NID:92622069; PIDN:AA8547
 A/Experimental source: strain Delta H
 C/Genetics:
 A/Genes: MTH978
 C/Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 31.0%; Score 741.5; DB 2; Length 455;
 Best Local Similarity 38.2%; Pred. No. 9,6e-40;
 Matches 173; Conservative 83; Mismatches 186; Indels 11; Gaps 5;
 Qy 9 VNGEMKSVNQIILSPIDSSLGFPVPMTRREVDHAKAGREALPMAALTYVERAQL 68
 Db 5 IDEEGVSGEKIFVTRNPFNGDEVDRVLAGRSVDERAIRAHRARADAMADLSARKISEKL 64
 Qy 69 HKRAADIIRKKEIATVLAKEISKAYNAVTEVTRADLIRYAEBSIRLSTADSGK 128
 Db 65 YDVADELKTELDEFARILITESGKPIRFSRDEKRSVETARLSAEEGR--YGSITPM 121
 Qy 129 DASTGKH--LAVIRROPVGIIVLAIAPNYPVNLSGSKIAIPALIGNVVMKPPQGSVSG 186
 Db 122 DADIGKGLGTFVRIPLGVVAAITPNNYPLNLAHKVGPALAAQTITIKPSLEAPLSA 181
 Qy 187 LVYAKPAEAGLPAGVNTTITGSGEIGDYIVHEEVPNFIPGSGTVGQRIQGLAGMRP 246
 Db 182 LKIAMLISE-HPPAGAVNAVATGSGSEVGVIIIDSPVDKITFGSVEVGRIYARASMK 240
 Qy 247 IMELGKDGKGIYVLAADLNNAAKOIVAGAYDYGORCTAIKRVLYVEEVDADSLAKISE 306
 Db 241 ITSELGNDPLIAMDADIDSAVEAIVKSYLYSGVCIAVKKAIYHEDVADFPADKLVN 300
 Qy 307 NVAKLVGDPEDNAT-VTPVIDNSADFIESTLVVDAROKAKELNEFKRGRLITPGLFD 365
 Db 301 ITGSLRAGDMDVTRVDVGLINDDAIEVERVIGAAVEDGABELLCGSGSRGNFVEPTVLD 360

Qy 366 HTLDMKAMEEPFGPIPLIRVYKDAEVAIAJANKSDFGLOSSVTRDFOKAFDIANKLE 425
 Db 361 HVPQKHEVIERETFGVSVIIRNENADEAIRLANGCYVALQAGVFENTIRALMARBEIE 420
 Qy 426 VGTVINNKTRGPNFPLGLKSGAGVQIGRI 458
 Db 421 AGTVLVNKSSTFRVDMFP-----GDGAVBEMVR 449

RESULT 15

A83033
 Probable aldehyde dehydrogenase PA4899 [imported] - Pseudomonas aeruginosa (strain PA01)
 C/Species: Pseudomonas aeruginosa
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 A/Accession: A83033
 R/Stover, C.K.; Pham, X.Q.; Ertwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A/Reference number: A82950; MUID:20437337; PMID:10984043
 A/Accession: A83033
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-489 <STO>
 A/Cross-references: UNIPROT:Q9HUR4; GB:AE004903; GB:AE004091; NID:99951173; PIDN:AG0828
 A/Experimental source: strain PA01
 C/Genetics:
 A/Genes: PA4899
 C/Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 31.0%; Score 741.5; DB 2; Length 489;
 Best Local Similarity 37.0%; Pred. No. 1.1e-39;
 Matches 173; Conservative 87; Mismatches 185; Indels 23; Gaps 7;

Qy 11 GEMK--SSVNOIILSPIDSSLGFPVPMTRREVDHAKAGREALPMAALTYVERAQL 68
 Db 16 GEMHGRAGRRLVNSPFDGSLLEIQADRDLDAVYKAAEVQPMALGPSARAVY 75
 Qy 69 HKRAADIIRKKEIATVLAKEISKAYNAVTEVTRADLIRYAEBSIRLSTSA---DB 124
 Db 76 YKAVEVFDHRHBEIVDMIRSEGS-----TRLKAEIEMGABARITLESASPARVH 126
 Qy 125 GGMKASTGKIAVIRROPVGIIVLAIAPNYPVNLSGSKIAIPALIGNVVMKPPQGSV 184
 Db 127 GRIVESDVGKESRYKSAIGVGVISPMNPFILHILORSIADPALAGNAVVKPASPDEV 186
 Qy 185 -SGVYAKAFAPAGLPAGVNTTITGSGEIGDYIVHEEVPNFIPGSGTVGQRIQGLAG 243
 Db 187 CGGLIARLFEERAGLPAGLFSVAVGSGSEIGDAFVHPVGLVTFGTGTPVGRNIGRIAS 246
 Qy 244 ---MRPIMLELGGKADGIVLADADLNNAAKOIVAGAYDYGORCTAIRVLYVEEVD 299
 Db 247 GGAHLKVALLEIGNSPFFVLGDADDEQAVNAVFGKFLHQOICAIIRIIVESDLVYA 306
 Qy 307 FAARFERYKGRVGPQPADTA-VGPIVNAQLBELLEKIRLARDEGAKPIYEGGVQ 365
 Db 358 LITPGLFDHVTLDKAMEEPFGPIPLIRVYKDAEVAIAJANKSDFGLOSSVTRDFOKA 417
 Qy 366 LLAIPHVGSVTATMEIARDEIFGPIGLRARDEAHLALANASEGYLSAASFRLERA 425
 Db 418 FDIANKLEVGTVINNKTRGPNFPLGLKSGAGVQIGRIYSIEAMT 465
 Qy 426 VAFARGLRAGMTHVNDIPVNDENAPFGGEKNSGLGRFPGDWAIBEFF 473

RESULT 16

A70318
 Aldehyde dehydrogenase - Aquifex aeolicus
 C/Species: Aquifex aeolicus
 C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

129 DASTGG--KLAIVIRPOVGIYLAIAPYNPVNLSSKIAIPALIGVNMFKPTQSSVG 186
132 ---PGERRVGVYFRPPIGIIYVATPNDPLNLVAHKIGPAAAGNSLIIKESPTPLA 188
187 LVLAFAAEAGLPAGVFNITTGSGEIGDIYVEHEBNFINFTGSPVQRIKLAGMP 246
189 LALVLAEEAGLPKIVQYITGSGVIGPLITKRVRLISFTGIGETSEKILARSAGVK 248
247 IMELGKADAGIYLAADADLNAAKQIVAGAYDSGRCIAIKRVLVVEEADLAEKISE 306
249 LAMEIGSNPTIYLAQDAEIMEAASCVSGAFCAAGNCIGVQIYVEGSEFNFSFGS 308
307 NNAKLSVDPFDPNAT--VTVIDNSADPESLVVDARQKAGELNFKDGLITGLD 365
309 QTKQLFSGKQSEETDIPMISEKEKRIERWEEKEKRGARVLTGRRRTGAFETPLVT 368
366 HTYLDKLMKEEPFGILPIIRYKDAEBAVALANKSDPGLQSSVFTRDFOKAFDIANKLE 425
369 NVSPSRLLAKEAFAPIVILIEGHSITETAIARANDVDPLQAGLFTNNILTAAPSALEKIE 428
426 VSTVHINNKTGRGPDNPFLGLKSGAGYOGIRYSIEAMTNVKSIVLDM 474
429 VGIWVNDSDVRIDAMPFGIKSGLGREGVRYAIEEMTEQKVAFHL 477

RESULT 19
F65045
succinate-semialdehyde dehydrogenase [NAD(P)] (EC 1.2.1.16) - Escherichia coli (strain K
C/Species: Escherichia coli
C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C/Accession: F65045
R./Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: F65045
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-482 <BLAT>
A/Cross-references: UNIPROT:P25526; GB:AB000351; GB:U00096; NID:gl789011; PIDN:AACT5708.
A/Experimental source: strain K-12, substrain MG1655
A/Genes: gahD
C/Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C/Keywords: oxidoreductase

Query Match 30.3%; Score 726; DB 2; Length 482;
Best Local Similarity 34.3%; Pred. No. 1e-38;
Matches 164; Conservative 92; Mismatches 206; Indels 16; Gaps 7;

6 QNVNNGEMKSSV--QIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPMAALTYVE 63
12 QALINGEMLDANNGEAIIDVTNPANGKLSVPMGADETRAAIDANRLDPAARALTAKE 71
64 RAQYIHKAAADIERDEKEIATVLAKEISKAYNASVTEVVRTDLIRYAAEGIRLSTSD 123
72 RATTILNNFMNMEHQDDLRLMTTLEQKPLAEAKGEISYAASFIMWPAEBEGRIYGDIT 131
124 EGGKMDASGKHLAVRRQPVGIYLAIAPYNPVNLSSKIAIPALIGVNMFKPTQSS 183
132 PGHQAD----KRLIYIKQPIGVTAAITPMNFPMAAITRKAGPALAAGCTMVLKPSQTP 186
184 VSGVLVLAFAAEAGLPAGVFNITTGSGEIGDIYVEHEBNFINFTGSPVQRIKLAG 243
187 FSLALAEALIRAGVAGFVNTVTSAGAVGNELTNPVLKLSFGSTIEIGQLMEQCA 246
244 --NRPIMLEIGSKDAGIYLAADADLNAAKQIVAGAYDSGRCIAIKRVLVVEEADLA 301
247 KQIKVSLIEGNAPIVFDADDLDKAVEGALASKFRNAGQTCVCANRLVYDGVYDRPA 306
302 EKISENVATLSVDPDPDN-ATVTPVIDNSADPESLVVDARQKAGEL---NEFRDGR 357

DB 307 EKLQAVSKLTHIGDGLDNVTITGPIIDEKAVAKVEHINDALEKARVYCGKAHERGDN 366
358 LITPLGFDHVTLDKLMKEEPFGILPIIRYKDAEBAVALANKSDPGLQSSVFTRDFOKA 417
367 FPGPTIIVDPANAKVSKKETEGPLAPLRFPEDEADVIAQANDTEGGLAAYFARLSRV 426
418 FPIANKLEVTGTHINNKKG-RGPDNFPFLGLKSGAGYOGIRYSIEAMTNVKSIVLDM 474
427 FVGEALREYGVGIN--TGIISNEVAPFGGIRASGLGREGSKYGLDYELIETKMGIGL 482

RESULT 20
139769
aldehyde dehydrogenase (EC 1.2.-.-) - Bacillus stearothermophilus
C/Species: Bacillus stearothermophilus
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: 139769
R./Imanaka, T.; Ohta, T.; Sakoda, H.; Wadhyastuti, N.; Matsuo, M.
J. Ferment. Bioeng. 76, 161-167, 1993
A/Title: Cloning, nucleotide sequence, and efficient expression of the gene coding for t
A/Reference number: 139769
A/Accession: 139769
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-488 <RES>
A/Cross-references: UNIPROT:P42329; GB:D13846; NID:g456298; PIDN:BA02975.1; PID:g456299
C/Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C/Keywords: oxidoreductase
F/49-309/Domain: aldehyde dehydrogenase homology <ALDD>
F/255,289/Active site: Glu, Cys #status predicted

Query Match 30.2%; Score 721.5; DB 1; Length 488;
Best Local Similarity 34.4%; Pred. No. 2e-38;
Matches 166; Conservative 104; Mismatches 192; Indels 21; Gaps 9;

3 KEYQNVNNGEMKSSV-NQIE-ILSPIDSSLGFPVPMTEBEVDHAKAGREALPMAAL 59
8 KTYFVYINGNWSVSNVNEPSINPANRHDIVGYQRTLEDVNEAVTANBAQTSWWR 67
60 TYVERAQYIHKAAADIERDEKEIATVLAKEISKAYNASVTEVVRTDLIRYAAEGIRLS 119
68 SEVGEVETVYKKAHLIEGGLDIAETMRMGKTLAEAKAETMRGHILIRYAGEAR-- 125
120 TSADEGKMDASTGKHLAVIRPOVGIYLAIAPYNPVNLSSKIAIPALIGVNMFKP 179
126 ---KIGDIVPSSDSGLLFTTRVPLGVGVISPMNFPVALPIYKMAPALVYGVTVLKPA 182
180 TCGSVGLVLAFAAEAGLPAGVFNITTGSGEIGDIYVEHEBNFINFTGSPVQRIK 239
183 SETAVTAAKVIECFHAGFPKGVNMVCGSGSVQGLANHPDIDGVFTGTSNTVQKQVG 242
240 KLAAGR--PIMLEIGKADAGIYLAADADLNAAKQIVAGAYDSGRCIAIKRVLVVEEVA 297
243 RAAPRGAKYQLEMKGKKNVYAKADADLDLAEVGTISGGLRSTGQKTAITSVFLIEREY 302
298 DELAKISENVAKLSVDPFDPNAT--VTVIDNSADPESLVVDARQKAGELNEFKR-- 354
303 EFPKAKLIERVQKLINGLDAETWNGPCASSQCHTVLSTYIEKGSBAGKLIYGNRCL 362
355 -----DGRILITGLFDHVTLDKLMKEEPFGILPIIRYKDAEBAVALANKSDPGLQSSV 409
363 EGBELNAGFVEFTIEDVDLQMTIAREELFPGVALLIQVDSIEBAIKLANDREYGLSASI 422
410 FTRDFOKAFDIANKLEVGTVHINNKTGRGPDNPFLGLK--GSGAGYOGIRYSIEAMTNV 467
423 YTKIGNALIEFKQIDIEAGLIKNAETAGVERPAPFGKMQSSSHSEGG-QAALIEFTSI 481
468 KSI 470
482 KTV 484

RESULT 21

A11546

succinate semialdehyde dehydrogenase homolog lin0913 [imported] - *Listeria monococca* (strain C) Species: *Listeria monococca*

C.Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C.Accession: A11546

R.Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karsc, U.

Science 294, 849-852, 2001

A.Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma Ok, C.; Schlueter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.

Article: Comparative Genomics of *Listeria* species.

A.Reference number: AB1077; MUID:21537279; PMID:11679669

A.Accession: A11546

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-488 <GLA>

A.Cross-references: UNIPROT:Q92DA3; GB:AL592022; PIDN:CAC96145.1; PID:g16413362; GSPDB:C

A.Experimental source: strain C1j11262

C.Genetics:

A.Gene: lin0913

C.Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 30.1%; Score 720; DB 2; Length 488;
 Best Local Similarity 33.6%; Pred. No. 2.5e-38;
 Matches 160; Conservative 100; Mismatches 198; Indels 18; Gaps 6;

QY 8 YNGBWSSVNO--IEILSPIDSSLGVPMTREVDHAKAGREALPAMALTYERA 65

Db 16 FINGMTIDGNKETDIYVNPANGVIAKIAQAGPSEKKAIAKADAFPMKAMELADRV 75

QY 66 QYLRADIIERDEKEIATVLAKEISKAYNAVTEVTRADLIRYAEGRILSTADEG 125

Db 76 KILHNIADMERADTLAKIMTLBOGKPLKESKGVLTGVENFRPAEERL-----YG 130

QY 126 GMDASTGHKLAIVIRRPVGIYLAIPYVNVLSGSKIAPALIGNVVMEKPTQSGVS 185

Db 131 ETIPAPNNHAF-IVKKQPIGVAAITPWNFGMVTTRKLAIPALATGNTIVLKPSGDTPLS 189

QY 166 GLVLAFAEAGLPAGVNTITGRSEITDYIVHEENFNFPGSTPVGQRIKLAG-- 243

Db 190 ALAIFEEAGLPKGVANIVGSSKEIGETLTSDVARKLFTGSTVGQTLFQOSADT 249

QY 244 MRPIMLEIGDGAGIVLADADLDNAKOIVAGAYVSGORCTAIRVLVEEVADELAEK 303

Db 250 LKKSLEIGAGAPFIVPADNDAEAVDILITAKFRNGVCVSPRIFPAKEIKKEFTKT 309

QY 304 ISENVAKLSVDPFDNATVTPVIDNSADFIESLVVDARQKAKELNEFKR-----DG 356

Db 310 LVAKVEKLVKVGNGLDVNVGPLIRDAIDKIDKQKNATDKAKAVLTGGERLTGSDYDKG 369

QY 357 RLITGLPDHTVLDKMLAEPPGPIPLIRYKDAEVAIAKNSDFGLOSVFTRDPQK 416

Db 370 NFYKTVLDNVTREMDIFYEETFGVPIPLIVFENBEALTEMANDSEFGLASFYTKDLAR 429

QY 417 AFDINKLEVGTVHNNKTGRGPDNPFPLGKSGAGVQIGRYISIEAMTNVSIYL 472

Db 430 VEKGAALLEYGVANGAEIAISNPER-PRGCVKHSFGRENHGYMEETIYQVAFINL 484

RESULT 22

A11188

succinate semialdehyde dehydrogenase homolog lmo0913 [imported] - *Listeria monocytogenes* C) Species: *Listeria monocytogenes*

C.Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C.Accession: A11188

R.Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karsc, U.

Science 294, 849-852, 2001

A.Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma Ok, C.; Schlueter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.

A>Title: Comparative genomics of *Listeria* species.

A.Reference number: AB1077; MUID:21537279; PMID:11679669

A.Accession: A11188

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-488 <GLA>

A.Cross-references: UNIPROT:Q8Y819; GB:NC_003210; PIDN:CAC98991.1; PID:g16410316; GSPDB:C

A.Experimental source: strain EGD-e

C.Genetics:

A.Gene: lmo0913

C.Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 30.0%; Score 718; DB 2; Length 488;
 Best Local Similarity 33.8%; Pred. No. 3.3e-38;
 Matches 161; Conservative 96; Mismatches 201; Indels 18; Gaps 6;

QY 8 YNGBWSSVNO--IEILSPIDSSLGVPMTREVDHAKAGREALPAMALTYERA 65

Db 16 FINGMTIDGNKETDIYVNPANGVIAKIAQAGPSEKKAIAKADAFPMKAMELADRV 75

QY 66 QYLRADIIERDEKEIATVLAKEISKAYNAVTEVTRADLIRYAEGRILSTADEG 125

Db 76 KILHNIADMERADTLAKIMTLBOGKPLKESKGVLTGVENFRPAEERL-----YG 130

QY 126 GMDASTGHKLAIVIRRPVGIYLAIPYVNVLSGSKIAPALIGNVVMEKPTQSGVS 185

Db 131 ETIPAPNNHAF-IVKKQPIGVAAITPWNFGMVTTRKLAIPALATGNTIVLKPSGDTPLS 189

QY 166 GLVLAFAEAGLPAGVNTITGRSEITDYIVHEENFNFPGSTPVGQRIKLAG-- 243

Db 190 ALAIFEEAGLPKGVANIVGSSKEIGETLTSDVARKLFTGSTVGQTLFQOSADT 249

QY 244 MRPIMLEIGDGAGIVLADADLDNAKOIVAGAYVSGORCTAIRVLVEEVADELAEK 303

Db 250 LKKSLEIGAGAPFIVPADNDAEAVDILITAKFRNGVCVSPRIFPAKEIKKEFTKT 309

QY 304 ISENVAKLSVDPFDNATVTPVIDNSADFIESLVVDARQKAKELNEFKR-----DG 356

Db 310 LVAKVEKLVKVGNGLDVNVGPLIRDAIDKIDKQKNATEKAKAVLTGGERLTGSDYDKG 369

QY 357 RLITGLPDHTVLDKMLAEPPGPIPLIRYKDAEVAIAKNSDFGLOSVFTRDPQK 416

Db 370 NFYKTVLDNVTREMDIFYEETFGVPIPLIVFENBEALTEMANDSEFGLASFYTKDLAR 429

QY 417 AFDINKLEVGTVHNNKTGRGPDNPFPLGKSGAGVQIGRYISIEAMTNVSIYL 472

Db 430 VEKGAALLEYGVANGAEIAISNPER-PRGCVKHSFGRENHGYMEETIYQVAFINL 484

RESULT 23

D85913

succinate-semialdehyde dehydrogenase [imported] - *Escherichia coli* (strain O157:H7, subserotype) Species: *Escherichia coli*

C.Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C.Accession: D85913

R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew, J.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Diallanta, E.; Potamoukis, K.; Apodaca, N.

Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A.Reference number: A85480; MUID:21074935; PMID:11206551

A.Accession: D85913

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-482 <STO>

A.Cross-references: UNIPROT:Q8X550; GB:AB005174; NID:g12517094; PIDN:AA657768.1; GSPDB:CN

A.Experimental source: strain O157:H7, substrain EDL933

C.Genetics:

A.Gene: gabD

C.Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 30.0%; Score 717; DB 2; Length 482;
 Best Local Similarity 33.9%; Pred. No. 3.8e-38;
 Matches 162; Conservative 92; Mismatches 208; Indels 16; Gaps 7;

Qy	301	AEKISENPAKSVEDPEPNAT-VTPVIDDNADPEFESLVADARQKAKELNEFKRDRLL	359
Dp	302	VEVEREKVSQLOGTGNPAPATLIGPLINYOIGRLOQLVKESVAQKVEGHVQULM	361
Qy	360	TRGEFDHTLMLKAMEEPGPIILIRVKOAEBAVAIANKSDPQOSVPTRPQKAFD	419
Dp	362	SPTLISEVTINMPAKKERIFGPILAIIRAKOBAEBAIAIANDSPYLSGISIFTSIHRGVQ	421
Qy	420	IANKLEVGTVINNKTRGPDNFPFLGKSGSAGQGI RYSIEAMTNKSI VLDMK	475
Dp	422	VAKQIDTMIHVNOQPVNAEBAHISFGGKSDGIGFGEGBWLDKFTYKWLSTOEK	477

RESULT 31

AE08339
succinate-semialdehyde dehydrogenase [NAD(P)] (EC 1.2.1.16) - *Salmonella enterica* subsp
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A>Note: This species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 15-Mar-2004
C:Accession: AE08339
R.Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A>Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AE08339
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-482 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05900.1; PID:GL6503874; GSPDB:GN00176
C:Genetics:
A:Gene: STY2911
C:Superfamily: NAD-dependent aldehyde dehydrogenase, aldehyde dehydrogenase homology
C:Keywords: oxidoreductase

[illegible]

RESULT 32
T44987
aldehyde dehydrogenase (EC 1.2.1.-) [imported] - Haloferax volcanii megaplasmid pHV3
C|Species: Haloferax volcanii
C|Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #ext_change 09-Jul-2004
C|Accession: T44987
R|Farahani, R.; Imbeault, J.C.; St Jean, A.; Chan, C.C.Y.; Allard, G.; Charlebois, R.L.
submitted to the EMBL data library, March 1997
A|Description: Hereditary instability of the megaplasmid pHV3, and filamentation in the
A|Reference number: Z22886
A|Accession: T44987
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-496 <FAR>
A|Cross-references: UNIPROT:O34184; EMBL:U95374; PIDD:AAE71806.1
A|Experimental source: strain DS2
C|Genetics:
A|Map position: megaplasmid pHV3
A|Genome: plasmid
A|Note: expressed during exponential growth
C|Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C|Keywords: oxidoreductase

Query Match 28.5%; Score 682; DB 2; Length 496;
Best Local Similarity 34.6%; Pred. No. 6; Ee 26;
Matches 166; Conservative 85; Mismatches 207; Indels 22; Gaps 8;

```

Db      15 YIDEMTESSESSEIAVEDPSTRETVAHVRGTEADVDAAEAAAEKQESMAAPARQ 74
Qy      66 QYLHKADIIERDKEELATLAKESKANNASTVEVTRADILRYAABESIRLSTADEG 125
Db      75 EVEQIOLANEXEYDEIIDLAEHVGGSRJMGFTSIQIASDHSAEATLPRR- ---RG 129
Qy      126 GKMDASTGCHKLAVIRQPVGIVLALAPYNPVNLGSKTLPALIGGVNVWFKEPQTQSVS 185
Db      130 EHVASNIPEGKENTVQKPKGVTVIISPMNFPMLMSRAVAPAPAAAGNAVVLKXSTNSPIT 189
Qy      186 -GLVLAFAPABAGLPACVFNTITGRGSEIGDYIVEHEEVPINFPGSTPVGORIGKLAM 244
Db      190 GGLLFAKLFEETDLPESGVNVVVTGRGSEIDRVAGHPESDVVAFPGSTGVKRGVSIAGE 249
Qy      245 R---PIMLEIGKQAGIVLADADDLNNAAKIVAGADYSGQRCTAIRVLVVEVADDELA 301
Db      250 NLAVPAM-ELGNNANAHVTEGADVDRADAAVGSPFHQGVCSINRHIVHESVYDEVV 308
Qy      302 EKISENVAKISVGPDPFNATVTV--PVIDNSGADFIJSLVVDARQGA-----KELNEF 352
Db      309 EKLTERAHNELPVGSAHEHDDTVVGPILIDESORDENMLGVTEVTAAGATLEFGGSTADLDGV 368
Qy      353 KRDRLLTPGLFDHVTLLDMKLAEBEPGPILPILIRVDAEBAVAALANKSDFGLQSSVTR 412
Db      369 D-DSLVVGVPYTLGSGVTNDMAAAREHNGPIAPVLPFSDVDVEAVELANATEYGLSGSVHAG 427
Qy      413 DFQAFADIANKLBYGVTHINNKTGRGPDNFPFLGLKSGSGAGVCGIRYSIEAMNVKYSIVL 472
Db      428 DLAGKEIARLMEYGNHVVNDQPIINDEAHVFPFGTSGAGVGTATNSDAFLDEIETDKWISL 487

```

RESULT 33

C69584
aldehyde dehydrogenase aldY - *Bacillus subtilis*
Cispecies: *Bacillus subtilis*
CDate: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
CAccession: C69584
R.Kunst, F. Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteiro
C.; Bron, S.; Brouillette, S.; Brutschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabre, C.; Ferrari, E.
Nature 390, 249-256, 1997
Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallero

QY 64 RAQVLAHKADIIERRDKESIALTVLAKELSKAINASTVEVRRADILRYAAGEGILSLTSAD 123

Db 72 RANKLRFRFMDLMTENQDDLARLMTIEQGRPLAEAAGELVAAYASFLEWGEBAKRIYGDPTI 131

QY 124 EGGKMDASTGHKLAVIRQPVGIVLAIAPVNVYVULSGSKTIALPALIGNVMPFPPTQGS 183

Db 132 PGHQPD-----KRIIVIKQPIGIVTRAITPPMFPFSMTIRKAGPINALAGCTMYLKPASGTP 186

QY 184 VSGVLAKAFABAGLPACVFNPTITGSGSEIGDYIYEHEEVNFIFETGSTPVGORIGK--L 241

Db 187 YSALATAELBARAGIPKGVFSVVTGSGAGEVGEILTSNPIYAKRLFPTGSTEIGRQIMABCA 246

QY 242 AGMRPIMELGSGKAGIVLADADDLNAKOIVAGAYIDYSGORCTRAIKVYLVEEYABELA 301

Db 247 QDIKVSEIETLGGNAPFIYFDPDADIDPAAVEGALISKYRNNGQTCVCANELLVYQDDGYDAFV 306

QY 302 EKISENVAKLTVGDPFED--NATVTPVIDDNSADFETSLVVDARQKAGKELNEFK---RDR 357

Db 307 DKLKRAAVAKKLINGLEAGVTTGPLIDAKAAVKEEHTADAVSKAKVVSQKPRHALGCT 366

QY 358 LITPGLFDHVLDMKLAMEBEPGPIPLTIRVDAEVAVALANKSDPGLQSSVFTFDFOKA 417

Db 367 FFEPIITVDVDPKNALVSKDETFEGPLAFVFRFDBAEFVILAMENDIEFGLASIFYRADLARV 426

QY 418 FDIANKLEVGVIHINNTGR--GPDNFPFLGLKGSAGVQGIIRYSIEAMTNVKSIVL 472

Db 427 FRVFEQLEFGWAGIN--TGLISNEVAPFGGKASGLGEGSKYGEIDLETIKYCL 480

```

RESULT 36
E75574
aldehyde dehydrogenase - Deinococcus radiodurans (strain R1)
C|Species: Deinococcus radiodurans
C|Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C|Accession: E75574
R|White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A|Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A|Reference number: A75250; MUID:20036896; PMID:10567266
A|Accession: E75574
A|Status: Preliminary
A|Molecule type: DNA
A|Residues: 1-524 <WHI>
A|Cross-reference: UNIPROT:O9RYT8; GB:AEO01863; GB:AEO01825; NID:g6460670; PIDDN:AAF124
A|Experimental source: strain R1
A|Genetics:
A|Gene: DRA0220
A|Map position: 2
C|Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match      27.9% Score 666.5 DB 2; Length 524;
Best local Similarity 34.4%; Pred. No. 6.9e-35;
Matches 166; Conservative 79; Mismatches 222; Indels 15; Gaps 7;

QY    1 LTKKYQNVTNGSEW--KSSVNQIEILSPIDDSLGFVPMTRBEEDVHAMKAGREALPAAVA 58
       ||:||:|||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
DB    27 LSRGIGHRTIGGEVMVHAHGSKTPDASHPVNDPLVVVAAGDASDIRAKKAHADAFQTIRE 86
       ||:|||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||

QY    59 LTVERAOYLHKADIIRDKEELATVLAKESIKAYNASVTEVPTADLRVYAEEGRRL 118
       ||:|||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
DB    87 VSGARRRKLHLKVADLIERADEIAVLSEVDTCQAIRFKMSAAGAENFRFYADR---- 142
       ||:|||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||

QY   119 STSDDEGGKMDASTGHKLAIVTRQPVGLVLAAPRNYPVNTSGSLIAPALLIGANNVMFXP 178
       ||:|||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
DB   143 APGADGGSLRPAGFINYSI--ROPISGPVGITPPMNTFMSTWKXIALPALAAGCTVVKXP 200
       ||:|||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||

QY   179 PTQGSVSGVLAKAPAEAGLPAGVFNTITGRSGEIGDYIVEHEEVNFINTGSTVPGQR 238
       ||:|||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
DB   201 AEMSVSAKLTLFEINDEAGLPBGVNNLVHGRFSAGKSLTEHPLYKAVAFLGETTTGSHI 260
       ||:|||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||

QY   239 GKLAG-MRPIMLELGKDAGIVLADADDLNNAKOIVAGAVDYSGORCTAIRVLVVEV 296

```

D6	261	MRGADTLTKRVHPELGGKRPVVYPFDADDDKDLDAVFMITSLNGERCSTSSRVLIDEGI	320
QY	297	ADLEAKRISENVAKUSVGDPEFDNAT-VYPIVDDNSADFIESLVYDARQKAKELNEKR-	354
D6	321	YDFYTRIRIAERANRIVGDEPLDPDTEVGPVLVPHRHFEEKVMSEYFECARQEGATTAAQGERV	380
QY	355	--DGRLLTFTGCLFDHVTLLDMKLABEERPFPIIPIIRVKDAEANAIAAKSPDQSSVFTR	412
D6	381	GEAGNFVVRPTLFTAAANDMKIAQEEIFGCVTLAIPFKDEADALIELADVANYGLAGLAYMTN	440
QY	413	DPQAKADINKLVEGVVTHINNTKGGDGNPFILGKSGCAVQGIKRSIEMTNVKSIVL	472
D6	441	DLRRAIRFAHGLEAGMIWVNSENVHLPF-PFGGVKNSGIRDGDGYSFPEFYMETKNIAI	499
QY	473	DM 474	
D6	500	SL 501	

RESULT 37

S14629
aldenhyde dehydrogenase (NAD) [EC 1.2.1.3] 1, cytosolic - chicken
CjSpecies: Gallus gallus (chicken)
CjDate: 21-Nov-1993 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
CjAccession: S14629
R.Godbout, R.

submitted to the EMBL Data Library, April 1991

A>Description: High levels of aldehyde dehydrogenase transcripts in the undifferentiated
A.Reference number: S14629
A.Accession: S14629
A.Molecule type: mRNA
A.Residues: 1-509 <GOD>
A.Cross-references: UNIROT:P27463; EMBL:X58869; NID:g63032; PIDN:CAA1679.1; PTD:g63033
A.Experimental source: retina
CjComplex: homotetramer
CjFunction:
A.Description: catalyzes the oxidation of an aldehyde to an acid using NAD+ and water
A.Pathway: ethanol catabolism
A.Note: enzymes with this activity are involved in diverse metabolic pathways in various
CjSuperfamily: NAD-dependent aldehyde dehydrogenases; aldehyde dehydrogenase homologous
F:67-331/Domains: aldehyde dehydrogenase homologous <LDL>
F:277,311/Active site: Glu, Cys #status predicted
F:464/Binding site: NAD (Cys) #status predicted

Query Match 27.8%; Score 666; DB 1; Length 509;
Best Local Similarity 33.9%; Pred. No. 7.2e-35;
Matches 163; Conservative 88; Mismatches 204; Indels 26; Gaps 9;

Dy YVNGEWSSVN-OIEITSLPDSLGFEVPAMTREEDHMKKAREAL---PAWALTYY 62
::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 32 FINEEMDVSQKKFEYFNPANEEKICEVAEGDKADDKAVAKRKAFELSGFRIMDAS 91
63 ERAQLTKAADIIDREXBIATVLA----KEISKAYNAVTEVVRTADLRVAABEGRIL 118
ERGLINKMLADVDRDLTLATWEAIDGKLFSFAVLMDAGACIKT---IRYCAGMADI 148
Db 92 ERGLINKMLADVDRDLTLATWEAIDGKLFSFAVLMDAGACIKT---IRYCAGMADI 148
119 STSADDEGGKNDASTGHKLAVIROPVGIVIALAIAPYNTPVALSSGIAPALIGNVVMFKR 178
149 H-----GRVTPEMDGNFFETFRHEFPVGCQIIIPWFPLMTFWKIAPALCCCNNTVVVKR 202
Dy 179 PTGSVSGVLVAKFAFAEAG.PAGVFNTITRGSEIGDYVEHEBWNINFTGSTPGORI 238
203 AEQIPPLSATLVGSLIKKAEGPPGVNNIVPFGCPLAGALISHMDDIVNSTGSTEVGKLI 262
Dy 239 GKLAG--MRPIMLEGGKAGIYLADADADDNNAKOIVAGAYDYSQGCRTAIRVLVEE 295
263 KEAAGKNLKRVLTIELGKGSPNIIFADADDDEAAFHIGLPFHQQGCCIAAGRIFVEEP 322
Dy 296 VADELAEKISENNVAKLSVGP-FDMAVTPTVIDDNSADPFESLYVDAROKGA---ELNE 351
323 IYDEFVRRIERARKYTLLPGVOGQPIDKEOPFOKITLDIESGKRKGATLEGCGGP 382

Query Match **27.8%** Score 664.5; DB 2; Length 495;
Best Local Similarity 34.2%; Pred. No. 8.6e-35;
Matches 167; Conservative 98; Mismatches 198; Indels 25; Gaps 10;

Cy 1 LTRKYONYNGEW--KSSVNOIEILSPIDDSLGFPMTREBYVDHANKAGREALPAMWA 58
 :
Dd 9 IANVGLFINGEFVKQSSDETIETVPATGETLSHITRAXDKDVDAHVAAQEAFESMEL 68
 :

Cy 59 LVYVEAQLLHKRADIIERPKKEIATYLAKESIAAYNASVTEVTADILIRAAE-GIR 117
 :
Dd 69 TSKSEEAQMRLDIGDKIMAQCKIAMETLTNNGPIRET-----TALDIPRAAHFHYP 122
 :

Cy 118 LSTADEGGMADASTGHKLAVIRROPVGIVLAIPYNPVNLSGSKIAPALIGANNVMRK 177
 :
Dd 123 ASVILETBEGTVNDIDKDWTIVRHPIGVGAVAANPMLLAANKIKAPIAAGTTIVQ 182
 :

Cy 178 PPTQGSVSGIVLAKAPAEAGLPAGVFNTITRGSEIGDIYVEHEEVNFINTGSTPVQR 237
 :
Dd 183 PSSSTPLSLLEVAKIFQEV-LPGVVNILTKSGSBSGNALFNHDGDVKLSFTGSTDVGY 241
 :

Cy 238 IGKLAG--NRPMILELGGKQAGIVLADADDNAKOIYAGAVDYSGQCTAIKRVLYVE 295
 :
Dd 242 VAEAAKHVLVPATLELGGKSANIILLDDANTDLAVEGIOLGILFNQGEVCASGRLLVHER 301
 :

Cy 296 VADELAEKISENAKLSVGPDPDNAT-VTPVIDNSADPIESLVVDAROKAGEL---NE 351
 :
Dd 302 IYDQLVPRIOEARSNINRYGDRPODEATMGSTGQDQDKISYIDAAKESDQILAAGR 361
 :

Cy 352 FKRDGKLLTFGLFDHTVL-----DMKLAMEEPFGPIILPIIRVKDAEEVAVALANSDFGL 405
 :
Dd 362 LTENG--LRKGFFPEPLIAPDPNMHLAOEIEGFVTLVKXDQDOEADIDANDSEYDL 419
 :

Cy 406 QSSVFRPDFOKAEDIANLKEVGVTHINNKTGRGPDNPFPLGLKSGAGVQGIRYSIEAMT 465
 :
Dd 420 AGGVFQNTRLRALINIAKAVTRGITWINT-YNQVEBGAFFGGYKTSIGIRETYKGLASNQ 478
 :

Cy 466 NVKSIIVLD 473

RESULT 38

C89778
hypothetical protein alda [imported] - Staphylococcus aureus (strain N315)
C|Species: Staphylococcus aureus
C|Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C|Accession: C89778

R|Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hatcori, M.; Ogasawara, N.; Hayashi, H.; Hiratsutsu, K.
L|Date: 357, 1225-1240, 2001
A|Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A|Reference number: AB9758; PMID:21311952; PMID:11418146
A|Accession: C89778
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-495 <KUR>
A|Cross-references: UNIPROT:O99X54; GB:BA000018; PID:g13700083; PIDN:BAB41382.1; GSPDB:C
C|Genetics:
A|Gene: alda
C|Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Db 479 QVKN1YID 486

RESULT 39
AF3469
aldenhyde dehydrogenase (NAD) (EC 1.2.1.3) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AF3469
R:DeIvecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujter, C.; Los, T.; Ivanova, R.; Mazur, M.; Goleman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letete, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-477 <KUR>
A:Cross-references: UNIPROT:O8YEV7; GB:AE008917; PIDN:AAL52921.1; PID:g17983768; GSPDB
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11740
A:Map position: 1
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: oxidoreductase

Query Match 27.7%; Score 664; DB 2; Length 477;
Best local similarity 35.3%; Pred. No. 8, 8e-35;
Matches 170; Conservative 80; Mismatches 202; Indels 30; Gaps 12;

5 YQNVNNGEMKSSVNOIEILSPIDSS-LGFPVPMATREEVNDAKMGREALPMAALTYVE 63
Db HONLAGEWVG-DDIANINPSDNDVGTARATADTKAIAAKAFPMSSSGILE 62
4 HONLAGEWVG-DDIANINPSDNDVGTARATADTKAIAAKAFPMSSSGILE 62
QY PAQYIAKADIIERDKESIAIVLAKESKAINASTEVVTRTDLIRYAEBSGRISTAD 123
Db 63 RHAIIRKTADEILARKKEILGRILSREKTLVEGETIRASQIFDFAGECIRLA---- 118
124 EGGKKDASTGHKLAV-IRROPYIVLAIAPVNPVNLGSKTAPALIGSNVMPKPTQG 182
Db 119 --GEVLPARPGIIGVEVREPVGVGIIITPMFPALIPAMKTAIPALCTGNTIVFKPAELV 176
183 SVSGIVLAKAFEAAGLPAGVFNITIRGSEIDGYIVHEBEVNFIFGTSGTPVGORI--GK 240
Db 177 PGCSMAIDILHRAGLPAGVNLVWGKSSVVGQTIILDSADVAAAFPTGSGTKGVAAAS 236
QY 241 LAGMRPIMLEIGKQAGIVLADADIDNNAKQIVAGAYVDSGORCTAIRKVLVEEVADEL 300
Db 237 IEHNRFPQLEWGGKKPVPVLLDDADINVAVESVNSAPFSTGQCTAASRIIYTEGIHDKF 296
QY 301 AEKISENVAKLSVGGPFDNAT-VYFVIDD---NSADFPESLVVDAROKGAK-----EL 349
Db 297 VAAAEIKTKTVDNALKRPETHIGPVVDETQLQODMDYIEL---GRKEGALKRGERL 352
QY 350 NEFKKDGGLTPGLFDHTYLLDMKLAWEPPGFIPIIRKDAEFAVALANKSDFGLQSSV 409
Db 353 NR-ETPGYLLQPALFTEATNQMRISREEIFGVASVIVKQVEBELATANDSFGLSSEI 411
QY 410 PTRDFOKAFDIANKLEVSTVHINKTGRGPDNFPFLGLKGSAG--VQGIKYSIEAMTV 467
Db 412 CTTSLKATYHKRNSAEAMVWNLPTAGVDVHFVPGKRGKSSFGPREQG-RYAAREYTVV 470
QY 468 KS 469
Db 471 KT 472

RESULT 40
E90504
hypotheoretical protein gapv-3 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

Db 84 ERRLRYKLADLERLRLRLATMESMNGKLYNSAYLNDLACITKTLRYCAGMADKIO-- 141
Qy 122 ADEGKMDASTGHKLAVIRROPGIVLALAPYPPVNLSSGSKTAPALLIGNVMMFKPPTQ 181
Db 142 ----GRTIPIDGNFFTYTRHEPIGVCGQILPMNFPLMLKIKGPAISCGNTVVAPADQ 197
Qy 182 GSVSGVLAKAPFAEAGIPACVNTTIGRGEISIDYVEHEEVNFINTGSPVQRIKGL 241
Db 198 TPLTALHVASLIMEAGPPGVNIVPGYPTAGAAISSHMDIDKVAFTGSTEVGKLIKEX 257
Qy 242 AG---NRPTMLEGKGAGIVLADADLDNAKQIVAGAYDSGRCATIKRVLVVEEVD 298
Db 258 AGSNKRYTLELIGKSPCTVLADADLNAVEPAHHGVFHYGQCCIAISRIFVESITD 317
Qy 299 ELAEKISENVAKLSVGDPE- DNATVTPVIDNSADPISLIVDAROKAK---ELNEFKR 354
Db 318 EFRRSVERAKKYILNPLTPVGTQGPQIDKEGYDKLIDIESGKKEGKMLECGGPMGN 377
Qy 355 DGRLLTPGLFDHYTLDMKLAMEEPFGPIILITRYKQAEBAVAIAKNSDFGLQSSVFTTDF 414
Db 378 KGYVQPTVFSNTDEIRIKAEIIFGVQOIMKFKSLDVIDIKRANNTFYGLSAGVFTKDI 437
Qy 415 OKAFDLANKLEVGTGHIINNKTRGPDNFPPLGLKSGAGVQIRYSIEAMTNYKSIYL 472
Db 438 DKRATISSALQAGTAVW-NCYGVVASQCPFGGFKMSGKRGELGEYGFHEYTEVKTIV 494

RESULT 46

C95964
probable aldehyde dehydrogenase protein (EC 1.2.1.-) [imported] - Sinorhizobium meliloti
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: C95964
R.Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chalin, P.; Vorholter, F.J.; Hernan
proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A/Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A/Reference number: A95842; M0ID:21396508; PMID:11481431
A/Accession: C95964
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1485 <KUR>
A/Cross-references: UNIPROT:Q92UV7; GB:AL591985; PIDN:CAC49379.1; PID:G15140865; GSPDB:G
A/Experimental source: strain 1021, megaplasmid pSymB
R.Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federpiet, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Weiler, D.H.; Wong, K.; Yeh, K.
A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A/Reference number: A96039; M0ID:21368234; PMID:11474104
A/Contents: annotation
C/Genetics: SMD21539;
A/Genome: plasmid
C/Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C/Keywords: oxidoreductase

Query Match 27.4%; Score 656; DB 2; Length 485;
Best Local Similarity 36.1%; Pred. No. 2,9e-34;
Matches 171; Conservative 84; Mismatches 201; Indels 18; Gaps 8;
Qy 9 VNGEMKSVNQIEILSPIDSSLGFPVAMTRREBYDHAKGREALPAMALTYEERAOYL 68
Db 17 IADRVLDTDRVAKRFPMNDTVGVPA---GRAEHAKEAFALAAAYOPKLTYYEROKIL 73
Qy 69 HKAADITERDKKEIATVLAKEISKAVNASVTEVTRADILRYAAEGIRLSTGDEGKM 128
Db 74 LATAEALAKKEIISVITTELEGISKADSLYEVGRAFDVFTLAGQMCIR-----DDGEIF 128
Qy 129 DAS-TGHKLA---VIRROPYGLATAPYPPVNLSSGSKTAPALLIGNVMMFKPPTQGSV 184

Db 129 SCDLTPHGAKRIFTMRBEPLTAISAITPENNHLNVAHVAALATNNCVVVKTELTQM 188
Qy 185 SGLVLAKAPAEAGLPACVNTTIGRGEISIDYVEHEEVNFINTGSPVQRIKGLAQM 244
Db 189 TALLADLIYEAGLPPEMLSVTGWPADIGEMITTPHYDLVTFTGSPVVGKLIANAHY 248
Qy 245 RPTMLEGKGDAGV---LADADLNAAKOIYAGAVDSGRCATIKRVLVVEEVDLA 301
Db 249 KQVLELGGNDPLILINDSIDDLAADAALVAGATKNSGQCTAVAKRILCOESVADRV 308
Qy 302 EKISENVAKLSVGDPEDNAT-VTPVIDNSADPISLIVDAROKAKELNEFKRDRLLT 360
Db 309 PLVLEBAKRLRGDPEMDRSTDTGTYHEGAALLFEERVRAABEGADLIYHGRSGALLP 368
Qy 361 PGLFDHYTLDMKLAMEEPFGPIILITRYKQAEBA-VALAKNSDFGLQSSVFTTDFQKAD 419
Db 369 PLYVDRVPHQSDLVLEETFGPIILPIYKVPDDDATITLSNRYAFGLSSGVTCTNDYRMQK 428
Qy 420 IANKLEVGTGHIINNKTRGPDNFPPLGLKSGAGV-QGIRYSIEAMTNYKSIYL 472
Db 429 YIAGLKVGTVNIWEVGYRIEMSPFGIKDSGNGYKEGVIEAMKSFNTVKTSL 482

RESULT 47

AD3300
succinate-semialdehyde dehydrogenase [NAD(P)] (EC 1.2.1.16) [imported] - Brucella melitensis
C/Species: Brucella melitensis
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C/Accession: AD3300
R.Delvecchio, V.G.; Kapatal, V.; Redkar, R.U.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldman, E.; Selikov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen Brucella meliten
A/Reference number: AD3252; PMID:11756688
A/Accession: AD3300
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1487 <KUR>
A/Cross-references: UNIPROT:O8YI07; GB:AE008917; PIDN:AAL51567.1; PID:G17982287; GSPDB:A
A/Experimental source: strain 16M
C/Genetics: BMEI0386
A/Genome: BMEI0386
A/Map position: 1
C/Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C/Keywords: oxidoreductase

Query Match 27.4%; Score 655; DB 2; Length 487;
Best Local Similarity 35.3%; Pred. No. 3.4e-34;
Matches 169; Conservative 81; Mismatches 201; Indels 28; Gaps 11;
Qy 6 QNVVNGEMKSVN--QIEILSPIDSSLGFPVAMTRREBYDHAKGREALPAMALTYE 63
Db 13 QCLVNGRMWDADAGTTIKTNPADSGVIGTVELSATIKELADLSAKALSGMAANTAKE 72
Qy 64 RAQYTHKAADITERDKKEIATVLAKEISKAVNASVTEVTRADILRYAAEGIRLSTGAD 123
Db 73 RAGILRKWFDLITIANADALIMTSEQKPLAEARGEVLYAASFTEWFAKEKV----- 127
Qy 124 EGGKMDA-STGHKLAVIRROPGIVLALAPYPPVNLSSGSKTAPALLIGNVMMFKPPTQ 182
Db 128 YEDTTPAPONGRLTVI-RQPGVTRPATTPMFPAMITRKAPALAACTMIVRADLT 186
Qy 183 SVSGVLAKAPAEAGLPACVNTTIGRGEISIDYVEHEEVNFINTGSPVQRIKGLA 242
Db 187 PLTALALGILAKAGIPAGVLIYVGKAREIABLSNTYAKLSTGSTGTEGRLL--WA 244
Qy 243 GRP-----IMLELGGKAGIVLADADLNAAKOIYAGAVDSGRCATIKRVLVVEEVD 298
Db 245 OCAPITIKRISLRLGNAAPITVDDADLDAVGAWSKTRNAGQTCVCANRTYVVGQYVD 304
Qy 299 ELAEKISENVAKLSVGDPEP-NATVTPVIDNSADPISLIVDAROKAKELNEFKRDR 357
Db 305 KFAEKILAAKVKELKVGNGTEPGVITGPMIEEKALITKVKAHIEDAVSKAKLITGKEKELG 364

A: molecule type: DNA
A: Residues: 1-487 <MHI>
A: Cross-references: UNIPROT: Q9RZE6; GB: AEO01862; NID: 96460468; PIDD: AAF1229

A1:Reference number: A96039; MUID:21368234; PMID:11474104

C:Genetics:
A:Gene: gabD4
A:Genome: plasmid
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: oxidoreductase

Query Match 27.2%; Score 650; DB 2; Length 490;

Best Local Similarity 33.6%; Pred. No. 7.1e-34;
Matches 160; Conservative 84; Mismatches 214; Indels 18; Gaps 9;

```
QY 9 VNGEM---KSSVNOGIELSPIDSSLGFPVPMTEBEVDHAMKAGREALPAAALTYERA 65
DB 23 IGAELMLDRSDSGKTDPVSNPATGEVAILPDMRSRSTARAIDAAHAAQMAEKTKERA 82
QY 66 QYLHKAADIIEERDKEEIVTLAKEISKAYNASVTEVTRADLIRYAAEGIRLSTSDRG 125
DB 83 AVLRLYLVLVANADDAITILTEMCKPLTEAKGELLYGASVYEMFGEAKVYGGDTITG 142
QY 126 GMDASTGHKLAIVIRROPVGIYLAIAIPYVNVLSGSKIAPALLIGNVMFKPPTQGSYS 185
DB 143 HQPD-----KRITIVKQPIGVAAITPMPNFMNMLARKLAPAAAGCAVSKPALETPLS 197
QY 166 GLVLAKAFEAAGLPAGVFTTITGRGS-ELGDIYIYHEEYVNFNFTGSGTPVGQRIKLAGM 244
DB 198 ALALALLAERAGLPAGVFNVLSTDSAEVGEKMCANDKVRKLTFTGISTVWG-KILMROGA 256
QY 245 RPIM---LELGSGDAGIYVLADADLDNAKQIVAGAYDYGQRCTAIKRVLYVEEVADELA 301
DB 257 DOIIMLGLLELGNAPPIVDDADLDAAVGAAGAVAKTRNNQTCVCANRIFVQAGIYDAPA 316
QY 302 EKISENVAKLSVGDPFD-NATVTPVIDNSADFIESLVVDARQAK-ELINEFKDGRLL 359
DB 317 ARLTAKVSEMTIGDGEPEVDAGPLISEKALAKVEEHIRDAVTKGADVLGNGAGGLFF 376
QY 360 TGLGDLHTYLDMLKAMEEPFGPILPIIRKDAEBAVALNKSDFGIQQSVFTRDQKAD 419
DB 377 EPTVLTGAATMDKIKAGEETFGVAPLFEKTEEEVMSMANKTEFGIASYFYSKDVSVYR 436
QY 420 IANKLEVGTGVIHNNKTGR-GPDNFPPLGKSGAGVQGIIRYSIEAMTNKSVILDM 474
DB 437 VAALEFYGMVGIN--TGLISTEVAFPGVKQSGQREGSKTGLDVTETKICLST 490
```

RESULT 53

A55684
aldehyde dehydrogenase (NAD) (EC 1.2.1.3) 6 precursor, salivary - human
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1995 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
C:Accession: A55684
R:Hsu, L.C.; Chang, W.C.; Hiraoka, L.; Hsieh, C.L.
Genomics 24, 333-341, 1994
A:Title: Molecular cloning, genomic organization, and chromosomal localization of an ad
A:Reference number: A55684; MUID:95213025; PMID:7698756
A:Accession: A55684
A:Molecule type: mRNA; DNA
A:Residues: 1-512 <Hsu>
A:Cross-references: UNIPROT:P47895; GB:U07919; NID:9995897; PIDN:AA79036.1; PID:9544482
C:Comment: This isozyme is found at highest levels in saliva, stomach, and kidney and at
C:Genetics:
A:Gene: GDB:ALDH6
A:Cross-references: GDB:364103; OMIM:600463
A:Map position: 15q26-15q26
A:Introns: 33/3; 66/3; 115/3; 159/1; 179/3; 222/3; 260/3; 295/1; 356/3; 411/3; 464/2; 48
C:Function:
A:Description: catalyzes oxidation of an aldehyde to an acid using NAD+ and water
A:Note: enzymes with this activity are involved in diverse metabolic pathways in various
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: NAD; oxidoreductase; saliva
F:70-334/Domains: aldehyde dehydrogenase homology <ALDH>
F:280.314/Active site: Glu, Cys #status predicted
F:467/Binding site: NAD (Cys) #status predicted

Query Match 27.2%; Score 650; DB 1; Length 512;

Best Local Similarity 34.0%; Pred. No. 7.5e-34;
Matches 164; Conservative 85; Mismatches 208; Indels 26; Gaps 9;

```
QY 8 YVNGEKKSVN--QELISPIDSSLGFPVPMTEBEVDHAMKAREL---PAAALTY 62
DB 35 FINNEMHESKSGKFPATCPSTRBOICEBEGDPEVDKAAVAAOVAFQSGSPMRRLDAL 94
QY 63 BRAQYHKAADIIEERDKEEIVTLAKEISK----AYNASVTEVTRADLIRYAAEGIRL 118
DB 95 SRGRLLHDLADVERDRATLALETMDGKFLHAFIDLEGICRT--LAYFGMAKXI 151
QY 119 STSADGSKMASTGHKLAIVIRROPVGIYLAIAIPYVNVLSGSKIAPALLIGNVMFKP 178
DB 152 Q-----GKITPTDNNVCFTRHEPIGVCGAITPMPNFMNMLWKMLPALCCGTMVLKP 205
QY 179 PTQGSVSLVLAKAPAEAGLPAGVFTTITGSGELGDIYIYHEEYVNFNFTGSGTPVGQRI 238
DB 206 ABQTPVLTALVIGSLIKEGFPFGVNVIPGQPTVGAATISSHPQINKLAFSTGEVGLV 265
QY 239 GKLA---GMRPIMLELGSGDAGIYVLADADLDNAKQIVAGAYDYGQRCTAIKRVLYEE 295
DB 266 KEAASRSNLKRVTLLEGSKNPCIACADLDLVBCHAGVFFNQGQCTTAASRVFVEQ 325
QY 296 VADEIAEKISENVAKLSVGDPFDNATVT-PYIDNSADFIESLVVDARQAK--ELINE 351
DB 326 VYSEFVRSSVYAKKRPVGDPEVDTEGQPQIDQKQFKEILIESGKKEGAKJECGSA 385
QY 352 FKRODRLLTPGLFDHVTLDMLKAMEEPFGPILPIIRVDAAEVAIALNKSDFGIQQSVFT 411
DB 386 MEDKLFKFTVFSVTDNMRIAKEBEIFGPQPLKFEISEIEIKRAASTVYGLAAVFT 445
QY 412 RDPKAPDIANKLEVGTGVIHNNKTGRGPDNFPPLGKSGAGVQGIIRYSIEAMTNKSV 471
DB 446 KNLDKALALASLEGVTWICNALYA-QAPFGGFMISGGRBELGEALAEYIEVKIVT 504
QY 472 LDM 474
DB 505 IKL 507
```

RESULT 54

AD0143
betaine aldehyde dehydrogenase (EC 1.2.1.8) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD0143
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tiltball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Parraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB00001; MUID:21470413; PMID:11586360
A:Accession: AD0143
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <KUR>
A:Cross-references: UNIPROT:Q8ZGV9; GB:AL590842; PIDN:QAC90007.1; PID:915979229; GSPDB:G
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: oxidoreductase

Query Match 27.1%; Score 649.5; DB 2; Length 490;

Best Local Similarity 31.8%; Pred. No. 7.6e-34;
Matches 153; Conservative 100; Mismatches 207; Indels 21; Gaps 10;

```
QY 8 YVNGEKKSVN--NOIEILSPIDSSLGFPVPMTEBEVDHAMKAGREALPAAALTYERA 65
DB 10 YINGVYTDSTSGDPTDANPANNGECIADQAAANQDVDAVAAKQGPVWAAAMTAMERS 69
QY 66 QYLHKAADIIEERDKEEIVTLAKEISKAYNAS-VTEVTRADLIRYAAEGIRLSTSDG 124
DB 70 RILRAVDILDRNDLEALETADTGKPLSETRSDIVTGADVLEYVA--GL---IPALE 124
```

[illegible]

RESULT 55

probable aldehyde dehydrogenase [imported] - fission yeast (*Schizosaccharomyces pombe*)
C1Species: Schizosaccharomyces pombe
C1Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C1Accession: T50272
R1Hunt, C.; Aves, S.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
Submitted to the public domain

A:Accession: T50272
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-496 <HUN>
A:Cross-references: UNIPROT:Q9UEW9, EMBL:AL133522, PIDN:CAB63554.1, GSPDB:GN00066, SPDB:
A:Experimental source: strain 972h(-); cosmid C922
C:Genetics:
A:Gene: SPDB:SPAC922.07c
A:Map position: 1
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

	Query Match	27.1%	Score 649	DB 2:	Length 496
	Best Local Similarity	33.1%	Pred. No. 8.3e-34		
	Matches	156	Conservative	97	Mismatches 203
				Indels	16
				Gaps	7
QY	8	YNGEWMKSVNQIELISPIDDSSLIGFVPMPTREEVDVHAKAGREALPAWALTVERACY	67		
DB	24	YINGEMHKSAAETWEIVDPSIEEVIAKVYLAKGEKIDVAYKSAKEAPKWKVPGSRKEL	83		
QY	68	LHKADDIERDKEEATVLAKEISKANVASTV-EVVRTADLIRYAAEEBIRLSTADDEG	126		
DB	84	LMKLAELTEKADDTLAIEANDSGKPLVSNAGVDGTTALRLYCGAMDKIT	137		
QY	127	KMDASTGHLAVIRRPVGIVLAIAPYNYVNLGSKIAPALIGCNVNFKEPPTGSSVG	186		
DB	138	QVIRPGRPKLVYAKRTPIGVCGQIAPWNYPLMWAGKIPALALAAGCIIIKSAETPTSL	197		
QY	187	LVIAKAPAEAGLPAGVFNITITGSGSEIGYIYEHHEWVNFINTGSRPVGQRIGKLA--SM	244		
DB	198	LYFALVLEAGRPKEGVNIIISGLGITVAGSMYAKHPRIDIKIATPGTSIKGVITYQQLASNU	257		
QY	245	RPIIMELGSKAGIYLADADLDNNAKQIVAGYVDSGQRCTAIKRVLVVEEVADELAEKI	304		
DB	258	KAVITTECGSKSPFLVEEDADLDQAYKMAIGIMYSGQICTNSRIYVDSYVDKTEIEF	317		

[illegible]

RESULT 56

1-pyrroline-5-carboxylate dehydrogenase [imported] - Nostoc sp. (strain PCC 7120)
C1Species: Nostoc sp. PCC 7120
A1Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C1date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C1accession: AC1874
R1Name: T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriyuchika, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, E
Nakazaki, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriyuchika, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, E
DNA Res. 8, 205-213, 2001
A1Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium Anabaena
A1Reference number: AB1807; M0ID:21595285; PMID:11759640

A:Accession: AC1874
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-996 <KUR>
A:Cross-references: UNIPROT:G8Y2D7; GB:BA000019; PIDN:BA872498.1; PID:G17129885; CSPDB:(
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr0540

Query Match: 27.1%; Score 649; DB 2; Length 996;
C:Superfamily: bifunctional proline dehydrogenase/delta-1-pyrroline-5-carboxylate dehyd

Best Local Similarity 34.3%; Pred. No. 2.2e-33;
Matches 160; Conservative 87; Mismatches 199; Indels 20; Gaps 9;

1 LTKYONVNGEWKSSVNOIEILSPIDSS-IGFVPANTREEVDHMKAGREALPAMAL 59

Db 504 LGKTYLPLNGEIVQTAEVIDSVNPSNFSEVIGKVLISVEQAEQAMQAKAALFFGWRRT 563

60 TYVERAQYLKADIIERDKELATVLAKETSKAYNASTVEVTADLIRYAEGRIS 119

Db 564 SYKERGILRRAGDLMQGRRAELSNIVLEVGKPYKEADGEVSEAIIDFCRYADMERLH 623

QY 120 TSADEGGMDASTGKHLAVIRRPQGIVLAIAPYNNPVNLSSKIAPALIGNVMEKPP 179

Db 624 QGINY--DVDGETNRYI----YQPRGIWVVISPMWNPALACGMTVAALVTGNCITLLKPA 677

```

Qy      180  TQGSVGLTLAKAFPAEAGLPAGVNTTITGRGSEIGDIYVEHEVNFINFTGSTPGORI - 238
          :      :      :      :      :      :      :      :      :

```

Db 678 ETSSVTTAKLTTELVEAGIPKGVQYVPGKGSQVGAYLVSHDTHLIAFTGSQEVGCRIV 737

```
QY      239 GKLAKRP-----IMLEGGKAGIVLADADLDNNAKOIYAGAYDYSQQRCRAIKRVL 291
```

Db 738 AEAATLKPGQRHMKRVIAEMGGKNAIVDESADLDQAVGVQSAFGISGQKCSACSRV 797

```
QY      292 WEEBADELAERKISENVAKLSTVGD-PFDNAFTTPTVIDNSADPIESLW--DARQGAKE 348
```

DB 798 VWEATDAF IHRVEATKSLNIGELPSPQVGPVADANARDRIREYIEKGAESQVALE 857

09 LINEKUGRLTFCGLFDHVLIDMKLAWEBEPGFIPIRIVKVAEAVAIANKSDGLOSS 408

349

958 TSA-DNHCVEVGDVTFCSEVBDCNTAIOOCTEGDITAVIKACHTMONATANTDNATTCG 916

[illegible]

Db 917 LYSRTSHIOAOEEFEVGNLYINRNITGAVARDFPGGPKSGVG 962

A:Cross-references: GB:M38433; NID:g145223; PIDN:AAA2428.1; PID:g145224
A>Note: the authors translated the codon AGG for residue 313 as Ser
C:Genetics:
A:Gene: aldH
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: NAD; nucleotide binding; oxidoreductase; P-loop
F:49-56/Region: nucleotide-binding motif A (P-loop)
F:58-322/Domain: aldehyde dehydrogenase homology <ALDH>
F:267,302/Active site: Glu, Cys #status predicted

Query Match 27.0%; Score 647; DB 1; Length 495;
Best Local Similarity 33.3%; Pred. No. 1,1e-33;
Matches 164; Conservative 95; Mismatches 203; Indels 30; Gaps 11;

```
QY 1 LTKXYQNYVNGEMKSVNQ-TEILSPIDSSIGFVPAMTREVVDHAKMAGEALPA-W 56
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 16 LAIENRLPFINGEYTAANEETFEVDPTQAPLAKRGKSYDIDRAMAAGVEFGDW 75
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 57 AALTYERQYLAKADITERDKEIATYLAKESKAYASV-TEVVTATDLIRAAEEG 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 76 SLSSPAKRAVANKADLMBAHAEBLALETDTPKPIHSDRDDIPGAARIRWYAE- 133
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 116 IRLSTASDE-GGKMDASTGKLAVEROPVIVLAIPYNPVNLGSKTAPALIGANYV 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 134 -----AIDVYGEVATTSHEILAMTVREVGVIATVPNFPDLICWKLGPALAGNSV 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 175 MEKPEPTQGSVSGLVAKAPAEAGLPAGVENITITGKSEIGDIYVEHEEVNFINFTGSTPV 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169 ILKPEKPSLAIRLAGLAKENGLDGVLVNTGTGHEAGQALSRNNDIDALAFGSTTT 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 GQRIGKLAG--MRPIMELGSKDGIYLA-GLDMAKQIVAGAYDSGRCITAIKEV 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 249 GKQLKIDGDSNMKRVMEAGSKSANIYFADCPDLCQAASATPAGIFYNQGVCIAGTSL 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 221 LVVEEVADELAKISENNAKSVGDPEDNA-TVTPVIDNSADFIESTIVVDKOKAGKL 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 309 LIEESTIADEFTALLKQQAQNMOPGHLDPATTTGTLIDCAHDSVHSFIREGSKG---- 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 NEFKRDKG-----LLTPGLFDHVTLLDMELAMEBEPGPIITIRVDAEAAVALANKSPF 403
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 365 -QLLDGRNAGLAALIGPIIFVDVDPNASTLSREIFGPIVLVTRFTSEQALQLANDSQY 423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 404 GLQSSVYTRDFOKAFDIANKLEVGYTHINNTGKRGDPNPFPLGLKSGAGVGCIRYSITA 463
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 424 GIGAAVMTETDLSPRAHMSRRLKAGSVFVANNYND-GDMTVPFGGYKQSGNGRDKSLHLEK 482
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 464 MTNKSIVLDMK 475
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 483 FTLEKTIWISLE 494
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 60
D69764
benzaldehyde dehydrogenase homolog ycnH - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: D69764
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteri
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc
A.; Ehrlich, S.D.; Emmerston, P.T.; Entlian, K.D.; Errington, J.; Fabre, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
A:Authors: Schleich, S.; Roche, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
Akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumeirein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: D69764
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-462 <KUN>
A:Cross-references: UNIPROT:P94428; GB:Z99106; GB:AL009126; NID:g2632653; PIDN:CAB12199.
A:Experimental source: strain 168
C:Genetics:
A:Gene: ycnH
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.9%; Score 644.5; DB 2; Length 462;
Best Local Similarity 33.5%; Pred. No. 1,5e-33;
Matches 157; Conservative 84; Mismatches 208; Indels 19; Gaps 7;

```
QY 18 NQIEILSPIDSSIGFVPAMTREVVDHAKMAGEALPAWALTVERQYLKADITIR 77
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 3 DQLYVNPATGCEIKITIPQASATVEBAIERSHQAFKTSKTSANERTSLKMWELIVE 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 78 DKEIATYLAKESKAYASVTEVVTADLIRAAEEGRILSTADEGKMDASTGHKLA 137
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 63 HKEELADLITKENGKPYGEAEVLYGAGYLEMFAEAKRV-----YGRVPAFTGKRI 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 138 VIRROPVIVLAIPYNPVNLGSKTAPALIGANVMEKPEPTQGSVSGLVAKAPAEAG 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 118 VTRQPVGPVAAITPMNFPNMITKAPALAGCTFIKRAPDTPLSAYELAKLAYERG 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 198 LPAGVENITITGKSEIGDIYVEHEEVNFINFTGSTPVQQRIGKLAG--MRPIMELGKND 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 IPKDVLYVIGGGEIGLVNFTSPKIRKITFTGSPVGLKMKNSADVTKVHSMELGHA 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 256 AGTIVADLDLMAKQIVAGAYDSGRCITAIKRLVYVEVADLAKISENNAKLSVSD 315
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 238 PLIVEDADIDILAVQMAASKRNAGQTCVCANRLIHESIIDEFAAKLSEGVSKLVGN 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 316 PF-DNATVTPVIDNSADFIESTIVVDKOKAGKL--NEFKRDKG-----RLTPGLFDHVT 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 298 GLIEEVNVPPIINRGFEKITYSDIDAVEKAKVAGTGYDRNDKCGCFVNPVLTVDV 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 369 LDMKLAMEBEPGPIITIRVDAEAAVALANKSPFGLQSSVYTRDFOKAFDIANKLEVGT 428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 368 TSMNIMHEETFGFVAVITYFSDIDBALQLANDTPGLAAYFTENYRRGIRYSEMLEYGI 417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 429 VAINNTGKRGDP--NFPPLGLKSGAGVGCIRYSIEANTNVSIVLDM 474
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 418 IGMND---GGPFAVQAPPGAKMESGIGREGSSEGIPELTKYLSIGL 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 61
C69814
benzaldehyde dehydrogenase homolog yfMT - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: C69814
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteri
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc
A.; Ehrlich, S.D.; Emmerston, P.T.; Entlian, K.D.; Errington, J.; Fabre, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
A:Authors: Schleich, S.; Roche, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
Akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumeirein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: C69814
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-465 <KUN>

A:Cross-references: UNIPROT:O06478; GB:Z99107; GB:Z9108; GB:AL009126; NID:g263055; PIR:
A:Experimental source: strain 168
C:Genetics:
A:Gene: yfmr
C:Superfamily: NAD-dependent aldehyde dehydrogenase, aldehyde dehydrogenase homology

Query Match 26.9%; Score 643.5; DB 2; Length 485;
Best Local Similarity 33.5%; Pred. No. 1.8e-33;
Matches 158; Conservative 86; Mismatches 216; Indels 11; Gaps 5;

OY QNYVNGEKK--SSVAQIEILSPIPDDSLGFPAPMAREEVVDHAMKAGREALPAMALTATVE 63
DB :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
9 KQFIGCKRQEBSRPVLNKNPYTKTFTTRKAATADVDAIYAALAACKMDVANVPFE 68
OY PAQYLHKADIIRDKEEIATVLAKESIKAVNAVTEVPTADLIRVAEEGRISTSD 123
DB ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
124 EGKKNDASTGHKLATRQPGIVLATLPYYPNVLSKSSTAPLLIGNNVMFPPQQS 183
DB EGIPLSPITDGENELTVYPAGVGVISPPNFPEFLSKSVAPALGAANGVALPHETTP 183
184 V-SGLVLAKAEAAEAGLVGFVFNTLTGRSGSEICGYIVEHEBNFNINFGTSVQGRIOLA 242
DB ICGGTLLINKITEENMGIPAGLLNVVTDLAEIGDSVEHPERRIIISFTGSTVGSYIQLA 243
OY --GNRPIMLEIGKGDAGIVLDADIDLNAKOIVAGAVDYSGQRCTAIKRVLVBEVDEL 300
DB MKHFFKPRLLEGNSAFVILEPADIEYAVNNAVFERSPTHQCQICMSANRVLVHSIIDKF 303
OY AEKISENVAKLSVGPDPFNATYT--PVIDDNSADPIESLVVAROKGAELNEFRDRLL 359
DB LELYPAKVESLKVGPMDDTIIGPLINSRQTDMKTVEQAIBEGAVPVLGGFNGTIV 363
OY TGPLFDHYTLDMKLAMEEPFCGPILPIIRVKOAEEVALANKSDPFOGSSVTRPOFAFD 419
DB EPTIAKDVKPFMSIAKEELFGPVVSFMKPDSEDEVDIANETPRGLSAVITSIERGVA 423
OY TANKELEVGVTHINNKTGRGPNDFPLGLKGSGAGVGYIRSIEANTVYSI 470
DB FAKRIETGMIHVNDDTTINDERPVAAGEKOGGLGRLNEMWGLEBETTLLKIWI 474

RESULT 62
CB3675
glycine betaine aldehyde dehydrogenase gbaA [imported] - Bacillus halodurans (strain C-1-
glaucina betaine aldehyde dehydrogenase
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: CB3675
R:Takami, H.; Nakabeone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirata,
Nucleic Acids Res 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: AB3650; MUID:20512582; PMID:11058132
A:Accession: CB3675
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <STO>
A:Cross-references: UNIPROT:O9KC49; GB:AP001507; GB:BA000004; NID:gi10172612; PIDN:BAB039
A:Experimental source: strain C-125
C:Genetics:
A:Gene: gbsa
C:Superfamily: NAD-dependent aldehyde dehydrogenase, aldehyde dehydrogenase homology

Query Match 26.9%; Score 643.5; DB 2; Length 490;
Best Local Similarity 32.9%; Pred. No. 1.8e-33;
Matches 161; Conservative 89; Mismatches 200; Indels 39; Gaps 8;

OY YONTVNGEWKSYNQIEILSPIPDDSLGFPVPMTREEDVHDAMKAGREALP--ANAALTIV 62
DB :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
6 YRMVINGERVSDDEFETTFNPAGKEILTATVAKASREHAEKAVQARAHAFDHGRKKRYPVG 65
OY ERAQYLHKADIIRDKEEIATVLAKESIKAVNASVEVPTADLIRVAEEGRISTSTA 122
DB :||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:

Db 66 KRAVVLNQIAIMRERQELVEISVLSGAIISAQAQINQIAIEDEFFA----- 115

Qy 123 DEGGKMDASTGHR-----LAVIRROPVSVILAIAPYNPVNLGSGKIAIAPALIGN 172

Db 116 -----GAIVGHRGTAVNNVNGFNLTQKEPVGCAQIIPMNPLPMAAAKVAIPALAGC 169

Qy 173 VVMFKPPTQGSVSLVLAFAEAGLPAGVFNITITGRGSIIGYIYIEHEVNFINTTGST 232

Db 170 SVVVKPASTLPPIITAILLINEICHEAGVPEGVNVNVAAGSVVAGDYLVGHEBVNVVAETGST 229

Qy 233 PVGRIQKLAG--MRPIMLEIGGKDAGIVLADLDLNAAQIYAGAAYDSGQCTAIKRV 230

Db 230 PIGQIDIEKASHTLAKRTLELGGKSPVLVADDMEAANAAGSLFGIYFNTGSCSEARSL 289

Qy 291 LVVEVADELAEKISENVAKLSVGDPEPDNAT-VTPVIDNSADFIISLVVDARQKA--- 346

Db 290 FVHEIYDFIEBIEBTRAKTLQGDPPDKCTHVGSIISRQLEITDSYVNSABEGATIA 349

Qy 347 ---KEAL-EFKRDRLLTPGLPDHYTLDMLKAEERPGLPIPIIRKVDAAEVAIANKSD 402

Db 350 LGGERVEVEGENHWYEPFVITNVTPDMKAVOEIEFGPVVETQDEBEVIRKANDTK 409

Qy 403 FGLDSSVPTDRPOKAPFIANKLEVGTVHINNTKGRGDPNPFGLGSGAGVGIGIRYSIE 462

Db 410 FGLGSALVTWDQGRATVNAHQLEAGIVMNVNCPSPAPPGT-PRGGYKESGREGRELCVETLD 468

Qy 463 AMTNVKGIV 471

Db 469 LYMEDKSTL 477

RESULT 63

E90863
aldehyde dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05050863)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: E90863
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. G. Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shnagawa, H. DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome A:Reference number: A99629; MUID:21156231; PMID:11258196
A:Accession: E90863
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-495 <HAY>
A:Experimental source: UNIRROT:Q87G6; GB:BA000007; PIDN:BA835300.1; PID:g13361342; GSPDB:A:G:Genetics:
C:Gene: Ecol877
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.9%; Score 643; DB 2; Length 495;
Best Local Similarity 33.1%; Pred. No. 2e-33;
Matches 163; Conservative 96; Mismatches 203; Indels 30; Gaps 11;

Qy 1 LTKKEYQYVNVGEMKSVNQ--IETLSPIDDSIGFVPAMTRREVDHAKMGKREALPA--W 56

Db 16 LAIENRLPFINGEYIAAENETFEYVDPVQAPLANIARGSVUIDRAVSAKGVFERGDM 75

Qy 57 AALTVERAOYLKAKADIERDKKEINTVLAKETSKAYNASV-TEVVRTADLIRYAAEEG 115

Db 76 SLSPAPKRVKLVNKLADLIENAEELALLTETLDGPKIRHSLDDIPGAPARIRWYAE-- 133

Qy 116 IRLSTASDE-GGKMDASTGKLAIVIRQPIGIVLALAPYNPVNLGSGKIAIAPALIGNV 174

Db 134 -----AIDKYGVAVTTSHELTAMIVPEPGVIAAIVPMNPLILTCMKLGPPLAAGNSV 188

Qy 175 MFKEPTQGSVGLVLAFAEAGLPAGVFNITITGRGSEIGDYIVHEHEVNFINTTSTPV 234

Db 189 VLKSEKSPISAIRLACLAEAGLPDGVNLVVTGFEHGAQALSRANDIDALIFGTSTRT 248

Qy 235 GQRIGKLAG--MRPIMLEIGGKDAGIVLADA-DLDNAKQIYAGAAYDSGQCTAIKRV 230

Db	249	GKÖLLKMDGSDSNMGRVWLEAGCKSANITYALOCPLDQAASTAGCIPIFNQÖVCLIACTRL	308
Qy	291	LVEBEVADDELAEKTSENVAKLSVGDPEDNA-TTTPVTDNSADPIESLVDARÖKAKEL	349
Db	309	LLEBSIADEFLALIKÖQÖNNQPGHPILDPAITWGTLLDCAHADSVESTIÖGESKQ---	364
Qy	350	NEFRDRG-----LITPGLFDHYTLDMKLAWESEPFGEILPIRVKDAEEVAALANKSDF	403
Db	365	-ÖLLIDGRNAGLAVAIQPTIIVDVPNASTSREEIFGVLVVTPTSEÖALÖLANDSÖY	423
Qy	404	GLOSSVFPEDÖQKAPDLANKLEVGTVHHNNKTGSGPDNFPPLGIKSGSAGVQÖGRYSIEA	463
Db	424	GLGAAVWTRDLSRRHRMSRRLKAQSVFVNNTND-GDMTVPPGGYKÖSGNGDKSLHAUEK	482
Qy	464	MTNVKSIVLDMK	475
Db	483	FTBELKTIWISLE	494

RESULT 64
D85755
aldehyde dehydrogenase, prefers NADP over NAD [imported] - Escherichia coli (strain O157:H7)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85755
R:Penn, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Hiller, U.; Grobbeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nattie 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85755
A:Status: preliminary
A:Molecule type: DNA
A:Releides: 1-195 <STD>
A:Cross-references: UNIPROT:O8XG6; GB:AE005174; NID:q12515474; PIDN:AA656504.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: aldH
A:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query	March	26.9%	Score 643;	DB 2;	Length 495;
Beat	Local Similarity	33.1%	Pred. No. 2e-33;		
Matches	163;	Conservative	96;	Mismatches 203;	Indels 30; Gaps 11;
Qy	1	LTKEYQNVNNGEMKSSVQNC--TEILSPIDBSLGFVPAMTREEDYDHAMKAGREALP--W	56		
Db	16	LAIENRLPEINEYTHAANENETFEETVPDPTQAPLANINARGKSVDDIDRAVSARGVFEEBGDW	75		
Qy	57	AALTYVERAQTLYHKAAADIIEPDKSEIATVTLAKESIKAYNASV-TEVVRTADLITYYAEEG	115		
Db	76	SLSPARKKAVLANKIADLIEANAEELALETLIDGKPRHSLRDIIDPEAAAIITWYAE--	133		
Qy	116	IRLSTSADE-CGKDDASTGHKLAVIRROPYGIYLAIPYNPVNLSGSKTAPALIGSNV	174		
Db	134	---AIDKYVGEVATTSHELMAYREPVGIVAIYPMNFPLLITCKLGPALAAAGNSV	188		
Qy	175	MFKPPTGQSVSGVLAKAPAEAGLPAGVFNTITGRGSEIGDYIYEHEBVNFINEFTSTVP	234		
Db	189	VLKSEKSPISAIIRLAGLAKRAGLPDGVLANVVTGSGHEAGQALSRHNDIDAIATFTGSTRT	248		
Qy	235	GQRIGKAG---MRPIMLEIGCKRAGIYLA--LIDNAAKIIVAGANDYSGQRCTAIKRV	290		
Db	249	GKQLLKDGDSNMKRWLEAGGKSANIVFACCPDIQQQAASATPAAGIFPNQOVCYIAGTRL	308		
Qy	291	LVEVEVADDELAKTISENVAKLSVGPEDFNA-TVTPVIDDNGSADEFIESLIVDAROKGAKEL	349		
Db	309	LLBESIADEFIALIKQQAQNVQPGHPDLPATWTGTLIDCAHADSVHSPIGGSEKSG----	364		
Qy	350	NEFKRDKG-----LITPGLFDHTYTLDMKLAHEBPFEGILPIRYKDAEEVAALANSDF	403		
Db	365	-QLLLDGNAGLAVAIIGFTIIVDVPDNPASISREELFGVLAIVTPTSEQALQLOANSOY	423		

[illegible]

RESULT 65
 T36807
 Probable aldehyde dehydrogenase - *Streptomyces coelicolor*
 C/Species: *Streptomyces coelicolor*
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C/Accession: T36807
 R/Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, July 1999
 A/Reference number: Z21614
 A/Accession: T36807
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-462 <OLI>
 A/Cross-references: UNIPROT:Q95246; EMBL:AL096681; PTDN:CAB46804.1; GSPDB:GN00070; SCOEI
 A/Experimental source: strain A3(2)
 C/Genetics:
 A/Gene: SCOEDB:SC130A.27c
 C/Superfamily: NMD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

[illegible]

```

RESULT 66
A33176
NAD-dependent succinate aldehyde dehydrogenases atkX [imported] - Agrobacterium tumefaciens
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: A33176
Rades, D.W.; Seubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
Mager, G.; Gillet, W.; Grant, C.; Guentherer, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClellan,

```


RESULT 71

T31293
aldehyde dehydrogenase homolog - Sphingomonas aromaticivorans plasmid pNL1
C:Species: Sphingomonas aromaticivorans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T31293
R:Romano, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G
submitted to the EMBL Data Library, July 1998
A:Description: Complete sequence of a 184 kb catalytic plasmid from Sphingomonas aromati
A:Reference number: Z20992
A:Accession: T31293
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-479 <ROW>
A:Cross-references: UNIPROT:O86001; EMBL:AF079317; NID:g3378261; PID:g3378434; PIDN:AAD
C:Genetics:
A:Genome: plasmid pNL1
A:Note: nahf
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.6%; Score 637; DB 2; Length 479;

Best Local Similarity 33.8%; Pred. No. 4, 6e-33;

Matches 153; Conservative 75; Mismatches 217; Indels 8; Gaps 3;

QY 23 LSPIDDSIGFVPAMTRREVDHAKAGREALPAMALTYVERAQLYHKAADIIRDEKEI 82

DB 24 LNPVDSVYSRVPAGAEVDNNAVENAHOAYLOHRLPAAVEGWIKAALMERDTAF 83

QY 83 ATFLAKEISKAYNAVSTEVVTRADLIRYAAEGIRLSTADSGKMDASTGKLAVIRRO 142

DB 84 ADVLVDELSPYAKAGFEFRFAVSFLRAAIGVPRRI-----RGETIPSDTPGRFSMLIQ 138

QY 143 PGIYVLAIPVYVNVNLSSKIAPALIGNVVMFKRPTGSGVGLAFAFAAGIPAV 202

DB 139 PGGVAVGITPVPVPLIKGKQSAAMALATNAPVLLPSEAPMADLALMKRAGVPDGL 198

QY 203 FNTITRGSGEIDYIYEHEVNFINTGSPGORIGKLA--GMRPIMLELGGKDAIYL 260

DB 199 FNVVYNGAGNEIDVLTGHRKVASITTTGSSRYGKHAETIAARNLKTYTELGGKSLVLC 258

QY 261 ADADLDNNAKOIVAGAYDYSGORCTAIKVLVVEVADELAEKISENVAKLSVGPFDNA 320

DB 259 ADADLDKAVNALFSIFPMYOGVCMGASRIYVERSIFDQFTAFAPAAATGRANSGLDRDPT 318

QY 321 TWT-FVIDNSADFTIESLVVDAROKAKALNEFKRDKRLITGLPHVTLDMKLAMEEPF 379

DB 319 TMLGPISRRQDRVNRHIDARSKGAVALAGGEMSGNSCAATILSGVTAEMTVPEEETP 378

QY 380 GPILPIRVKDAEBAVALANKSDFGLOSVPFRDFQKADIANKLEVGTVHINNKTRGRP 439

DB 379 GVTSLFFPDTLEAELELNNTREYLSASIFTRDLKALAFORABAGVNHINAPTLHDE 438

QY 440 DNFPLGLKSGAGVQGIYSIEMNTVKSIVL 472

DB 439 PHVPFGTXYASGFGREGTEADLEIMTEWKMTI 471

RESULT 72

F98208

succinate-semialdehyde dehydrogenase PA0265 [imported] - Agrobacterium tumefaciens (stra

C:Species: Agrobacterium tumefaciens

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C:Accession: F98208

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wolliam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2333-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: F98208

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-484 <KUR>

A:Cross-references: UNIPROT:O8U852; GB:AE007870; PIDN:AAK89192.1; PID:g15159010; GSPDB:GR

C:Genetics:

A:Gene: AGR L 1228

A:Map position: linear chromosome

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.6%; Score 636; DB 2; Length 484;

Best Local Similarity 33.2%; Pred. No. 5, 4e-33;

Matches 157; Conservative 82; Mismatches 218; Indels 16; Gaps 7;

QY 9 VNGEKSSVNO--IETLSPIDDSIGFVPAMTRREVDHAKAGREALPAMALTYVERAQ 66

DB 15 VAGRWIGCDDEBTATIRNPATGKIVGVPELGATETQATTAAYAQAMARRTGERAA 74

QY 67 YLHKAAADIIRDEKEIALVLAKEISKAYNAVSTEVVTRADLIRYAAEGIRLSTADSG 126

DB 75 ILKANHRLVMEKNDLGMILTLBQKPLAEKAGIAYGASFLWFAEARRRINGETVPQH 134

QY 127 KMDASTGKHLAVIRROPVGIYLAIPVYVNVNLSSKIAPALIGNVVMFKRPTGSGVSG 186

DB 135 QAD-----KRLIVLRQPVGVAAITPPNFPNAMIITRKGPALAGCAVVLKAPDTPESA 189

QY 187 LVLAFAFADGLPAGVNTITRGSEIDYIYEHEVNFINTGSPVGORIGKLAG--M 244

DB 190 IALAILARAGLPELFSIVGPAAEIGVLTASPDILTLFTGSTRGEBHLYRCCAPTI 249

QY 245 PRIMELEGKQAGIYADADLDNNAKOIVAGAYDYSGORCTAIKVLVVEVADELAEKI 304

DB 250 KGLGLELGGNAPFIFENDADDAVEGALIAKFRNNGTCCANFLYQSGVYDFAKPF 309

QY 305 SENVAKLSVGPFDNAVIT-FVIDNSADFTIESLVVDAROKAKEL--NEFKRDKRLIT 360

DB 310 ARAVSGLVKNGLESGSLGPIIDGNVAKYBAHIALDSLGAQQAAGSSRYALGPNFE 369

QY 361 PELPHVTLDMKLAMEEPFPIPIRVKDAEBAVALANKSDFGLOSVPFRDFQKADI 420

DB 370 PTLIRNVATDAQVAEEFTGFLAPLFRFEDDEVDYIAQANNDFGLASVYFARDLSRVERV 429

QY 421 ANKLEVGTVHINNKTR--GPDNFPFLGLKSGAGVQGIYSIEMNTVKSIVL 472

DB 430 AEALSYGVNVN--TGAISTABAPFGVGMKSLGREGSRHGLEETELKYVCI 480

RESULT 73

AC3078

aldehyde dehydrogenase atck [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C:Accession: AC3078

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClellan,

A.; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krepsan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AC3078

A:Status: preliminary

A:Molecule type: DNA

A:Cross-references: UNIPROT:O8U852; GB:AE008689; PIDN:AAL45041.1; PID:g17742704; GSPDB:GR

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: atck

A:Map position: linear chromosome

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.6%; Score 636; DB 2; Length 484;

Best Local Similarity 33.2%; Pred. No. 5, 4e-33;

Matches 157; Conservative 82; Mismatches 218; Indels 16; Gaps 7;

```

QY 9 VNGEMKSVNO--IETLSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTYVERA 66
DB 15 VAGRMIGCDRETATIRNPAATGKIYQVVELGATEQOAAITAAVIAQKMAARTAGERAA 74
QY 67 YLHKADIIERDEKEIATVLAKEISKAVNASVTEVVRTADLIRYAAEGIRLSTADSG 126
DB 75 ILKAMHRLVEMNRDDGLMILTLEQKPLAEAGELIYGASFIEMFEEBARRINGEVP 134
QY 127 KNDASTGHLAVIRROPVGIYLAIAFYNTPNLGSGLAPALIGNVVMFKPTQSGV 186
DB 135 QAD-----KRIIVLRPQGVVAATIPWNPNNMITKRVGPALAGCAVILKPAQPF 189
QY 189 LVLAFAFAGLPAQGFNTITGRSGEIRGYIYVHEEVENFINFTGSPVQRIKLAG--M 244
DB 190 IMLALIAERAGLPLEFISVITGPBAEIGVLTASPIRLITFTGSTRTEBHLIRQCAP 249
QY 245 RPIMLELGGKAGIYVADADLDNAAQIYVAGADYSGQRCTAIKRLVVEEVADEAEK 304
DB 250 KTLGIELGGMNAPFIVENDADLDAAVAGALIAKFRNNGQCTVCANRLYVSGYDAFA 309
QY 305 SENYAKLSVGDPPDNATVT-PVIDNSADFIESLVVDARQKAKEL--NEKRDGRLL 360
DB 310 ARAVSGIKVNGLEBESSIIGPLIDGNAVAKVEAHIDALSKGQIAAGSRYALGNGFE 369
QY 361 PGLFPHVTLDMKLAEEPPGILPIIRVKAEEVAVALANKSDPGLQSSVPTDFOKAF 420
DB 370 PTLIRNVTAQMOVAARETGPPLPFRFDEEDVIAQANNTPGLASIFYARDLSRVFR 429
QY 421 ANKLEGTGVIHNNKTR--GPDNPFELGKSGAGVQIGIRYSIEAMTVKSIYL 472
DB 430 ABALEVGWGVN--TGAISTAEFPQGVVMGSLGRGSHGIEHYELKYYVCI 480

RESULT 74
S74224
aldehyde dehydrogenase (NAD) (EC 1.2.1.3) 2 - mouse
N.Alternate names: retinaldehyde-specific dehydrogenase
C.Species: Mus musculus (house mouse)
C.Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 09-Jul-2004
C.Accession: S74224
R.Zhao, D.; McCaffery, P.; Ivins, K.J.; Neve, R.L.; Hogan, P.; Chin, W.W.; Draeger, U.C.
Eur. J. Biochem. 240, 15-22, 1996
A.Title: Molecular identification of a major retinoic-acid-synthesizing enzyme, a retinal
A.Reference number: S74224; MUID:96390857; PMID:8797830
A.Accession: S74224
A.Molecule type: mRNA
A.Residues: 1-499 <ZHA>
A.Cross-references: UNIPROT:Q62148; EMBL:X89273; NID:g1430868; PIDN:CA67666.1; PID:g143
A.Experimental source: strain C3H/He, Cell type embryonal carcinoma; cell line P19 terat
C.Gene: RALDH-2
C.Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C.Keywords: NAD; oxidoreductase
F.57-321/Domain: aldehyde dehydrogenase homology <ALDH>
F.193-273/Domain: NAD binding #status predicted <NAD>
F.1267,301/Active site: Glu, Cys #status predicted
F.454/Binding site: NAD (Cys) #status predicted

Query Match 26.5%; Score 635; DB 2; Length 499;
Best Local Similarity 33.5%; Pred. No. 6, 5e-33;
Matches 161; Conservative 84; Mismatches 210; Indels 26; Gaps 9;
QY 8 VYNGEMKSVN--QIEILSPIDSSLGFPVPMTRREVDHAKAGREAL--PAMALTYV 62
DB 22 FINNEQNSGSGVFPVCPATGEQVCEVOEADKVIDRAVQAARLAFSLGSVWRMDAS 81
QY 63 BEAQYVLRKADIIERDEKEIATVLA-----KEISKAVNASVTEVVRTADLIRYAAEGIRL 118
DB 82 EGRRLDLDKADLVERBATLATWESLNGKPLQAVYIDLDQGIK--LRYAGWADKI 138
QY 119 STSADGEGHMDASTGHLAVIRROPVGIYLAIAFYNTPNLGSGLAPALIGNVVMFKP 178
DB 139 H-----GMTIPVDGDFYFTTRHEPFGVCGQIIPNPFLLMFWKIAAPLCCGNTV 192

```

```

QY 179 PTQSGVSGIYLAKAPAEAGLPAGVENTITGRSEIIGDYIYVHEEVENFINFTGSPVQRI 238
DB 193 AEOTPLSALYMALIKAGFPFGVNNILPGYPTAGAAASHIGIDKIAFTSTVEGKI 252
QY 239 GLIAG--MRPIMLELGGKAGIYVADADLDNAAQIYVAGADYSGQRCTAIKRLVVE 295
DB 253 QEAAGRSNLKRYTLELGGKSPNIIIFADADLDVADQAGVFPNGQCTAASRIIVE 312
QY 296 VADELAEKISENVAKLSVGDPPDNATVT-PVIDNSADFIESLVVDARQKAKELNEFK- 353
DB 313 IYEEFVKSVERAKRIVSGFPDPTTEGQPDQKQYNNVELLIQSGVLEKALECGKG 372
QY 354 --RDGRLLTPGLFPHVTLDMKLAEEPPGILPIIRVKAEEVAVALANKSDPGLQSSVPT 411
DB 373 LGRKGFPEFPVFSVNTDMDRIAKEIEFGVQEIIRFKTMDVEIRANNSDFGLVAAP 432
QY 412 RDPKAPDIANKLEGTGVIHNNKTRGPDNPFELGKSGAGVQIGIRYSIEAMTVKSIYL 471
DB 433 NDINKALVSSAMQAGTWINCYNALMAOS-PFGGFKMSGNGRNGEFGIREYSEKIVT 491
QY 472 L 472
DB 492 V 492

RESULT 75
H98159
Hypothetical protein AGR_L1010GM [imported] - Agrobacterium tumefaciens (strain C58, Ce
C.Species: Agrobacterium tumefaciens
C.Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C.Accession: H98159
R.Goodner, B.; Hinkle, G.; Galtung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A.Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A.Reference number: A97359; MUID:21608551; PMID:11743194
A.Accession: H98159
A.Molecule type: preliminary
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-483 <KUR>
A.Cross-references: UNIPROT:Q8U711; GB:AE007870; PIDN:AAK88802.1; PID:g15158555; GSPDB:
C.Gene: AGR_L1010GM
A.Map position: linear chromosome
C.Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.4%; Score 632.5; DB 2; Length 483;
Best Local Similarity 34.0%; Pred. No. 9e-33;
Matches 161; Conservative 89; Mismatches 210; Indels 13; Gaps 5;
QY 8 VYNGEMKSVN--QIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTYVERA 65
DB 7 FIGDGHCAANRKRFRNNPIFGDAVTSAAASVEDATNADAAAAPFKWATTPGERR 66
QY 66 QYIHRADIIERDEKEIATVLAKEISKAVNASVTEVVRTADLIRYAAEGIRLSTADSG 125
DB 67 RHLTAAEALRAAGQIIOAMMEIEGATEAMGFVNLASDMLVEAAS-----LTHING 121
QY 126 GMDASTGHLAVIRROPVGIYLAIAFYNTPNLGSGLAPALIGNVVMFKPTQSGV 185
DB 122 EVIPSNRPETMAAROPAGVLSMAPMNAVPLIGVRSATPLACGNTVMTSELCPRT 181
QY 186 GLVLAFAEAGLPAGVENTITGR--GSEIGDYIYVHEEVENFINFTGSPVQRIKGLA 242
DB 182 HALIIEAVASAGLPQGVNAISNAPDDAKIVETILAHPAVARVNFQSTGRVIAEFA 241
QY 243 G--MRPIMLELGGKAGIYVADADLDNAAQIYVAGADYSGQRCTAIKRLVVEEVADEL 300
DB 242 GRVLRPALLELGGKAPFIVLDADIDAAVAAAFAFAYNMGGQICMSTERIIVLSDIDF 301
QY 301 AEKISENVAKLSVGDPPDNAT-VTVIDNSADFIESLVVDARQKAKELNEFKDGRLL 359

```

Db	302	VEAFAKKAASLTAGDPREKGTPLGSLVSAEASRIVLVDDAVSKARRVAGGGDTML	361
Qy	360	TPGLFDHYTLDMKCLAMEEPFGPILPIIRYKDAEEAVATANKSDPGLQSSVFTRDFOKAPD	419
Db	362	DAIAVDGVTTPARKLYSESPFVPSIIRAGSIDEAVSIAMESEFGLSAAVFGDDQARALS	421
Qy	420	IANKLEVGTVHINNKTGRGPDNFPFLGLKSGAGVQGIIRYSIEAMTNVKSIVL	472
Db	422	VAARIESGICHVNGPTVHDEAQLPFGGVKASGVGRFGGTAGIAEFTELRWTL	474

Search completed: August 28, 2005, 10:28:32
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2005, 10:05:09 ; Search time 176 Seconds

(without alignments)
1382.032 Million cell updates/sec

Title: US-09-868-195-12

Perfect score: 2393
Sequence: 1 LRKEIONVNGEKSVNOI.....GIRYSIEANTNYSIVLDMK 475

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 80 summaries

Database : uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2391	99.9	475	2	Q8RP84 streptococ
2	2382	99.5	475	2	Q8E0B1 streptococ
3	2379	99.4	475	2	Q8E5Y8 streptococ
4	1812	75.7	474	2	Q970T0 streptococ
5	1802	75.3	474	2	Q8DPS7 streptococ
6	1700	71.0	475	1	GAPV STRMU
7	1645	68.7	475	2	Q8P0F5 streptococ
8	1639	68.5	475	2	Q99267 streptococ
9	1618	67.6	467	2	Q8K707 streptococ
10	1383	57.8	482	2	Q97D25 streptococ
11	1345	56.2	479	2	Q73CX4 streptococ
12	1344	56.2	479	2	Q6HMX0 streptococ
13	1341	56.0	479	2	Q63FP8 streptococ
14	1338	55.9	479	2	Q81UL6 streptococ
15	1335	55.8	477	2	Q72Z26 streptococ
16	1333	55.7	479	2	Q81HE6 streptococ
17	1333	55.3	482	2	Q8XHP4 streptococ
18	1244	52.0	481	2	Q9KXN0 streptococ
19	1171.5	49.0	496	2	Q9SNX8 streptococ
20	1168.5	48.8	496	2	Q9ZUG8 streptococ
21	1164.5	48.7	498	1	GAPV MAIZE
22	1161.5	48.5	496	1	GAPV_PEA
23	1158.5	48.4	499	2	Q6ZNG0 streptococ
24	1155.5	48.3	496	1	GAPV NICPL
25	1145.5	47.9	507	2	Q8S4Y9 streptococ
26	1138.5	47.6	508	2	Q84PC4 streptococ
27	1120.5	46.8	471	2	Q6MKT9 streptococ
28	1077.5	45.0	472	2	Q6FLK7 streptococ
29	1047.5	43.8	478	2	Q8EVT9 streptococ
30	1047.5	43.8	478	2	Q8EVT9 streptococ
31	1027.5	42.9	496	2	Q8LKE1 streptococ

32	1026.5	42.9	496	2	Q8L5J9 streptococ
33	944	39.4	525	2	Q7NBX5 streptococ
34	868.5	36.3	475	2	Q9PQC9 streptococ
35	810.5	33.9	475	2	Q7A5E6 streptococ
36	790	33.0	478	2	Q75TD2 streptococ
37	786	32.8	478	2	Q8ZVJ1 streptococ
38	763	31.9	488	1	DHAL_BACST
39	759.5	31.7	463	1	YELL_METUA
40	752	31.4	470	2	Q97XAS streptococ
41	751	31.4	468	2	Q976X5 streptococ
42	750.5	31.4	486	2	Q8EMK2 streptococ
43	750	31.3	470	2	Q97X89 streptococ
44	742	31.0	497	2	Q8CV96 streptococ
45	741.5	31.0	455	2	Q27059 streptococ
46	741.5	31.0	484	2	Q8NMB0 streptococ
47	741.5	31.0	489	2	Q9HUR4 streptococ
48	741.5	31.0	496	2	Q6M2H6 streptococ
49	739.5	30.9	488	2	Q65NXX streptococ
50	739	30.9	506	2	Q8U2S5 streptococ
51	737	30.8	476	2	Q8E573 streptococ
52	736.5	30.8	491	2	Q59702 streptococ
53	736	30.8	474	2	Q81QX6 streptococ
54	736	30.8	474	2	Q6HJ76 streptococ
55	735	30.7	474	2	Q738S1 streptococ
56	735	30.7	482	2	Q9KC36 streptococ
57	734	30.7	471	2	Q8TWC7 streptococ
58	732	30.6	474	2	Q63BQ7 streptococ
59	726	30.3	482	1	GABD_BCOLI
60	726	30.3	488	2	Q721Q0 streptococ
61	725	30.3	499	2	Q8RFB1 streptococ
62	722.5	30.2	483	2	Q73EK2 streptococ
63	721.5	30.2	483	2	Q81ZE2 streptococ
64	721.5	30.2	483	2	Q6HP75 streptococ
65	721.5	30.2	488	1	DHAL_BACST
66	720	30.1	488	2	Q92DA3 streptococ
67	718.5	30.0	483	2	Q63GQ3 streptococ
68	718	30.0	488	2	Q8Y819 streptococ
69	717	30.0	482	2	Q8X950 streptococ
70	716.5	29.9	493	2	Q9HK01 streptococ
71	716	29.9	474	2	Q81DV8 streptococ
72	713.5	29.8	480	2	Q9H0Z2 streptococ
73	713	29.8	501	2	Q57693 streptococ
74	711.5	29.7	472	2	Q8EMH9 streptococ
75	711.5	29.7	483	2	Q81IM7 streptococ
76	711	29.7	475	2	Q8BLI8 streptococ
77	710.5	29.7	489	2	Q9HMT6 streptococ
78	710.5	29.7	498	2	Q8XOP9 streptococ
79	710	29.7	486	2	Q7VSP3 streptococ
80	710	29.7	486	2	Q7W373 streptococ

ALIGNMENTS

RESULT 1
Q8RP84 PRELIMINARY; PRT; 475 AA.

AC Q8RP84; 01-JUN-2002 (TREMBLER, 21, Last sequence update)
DT 01-JUN-2002 (TREMBLER, 21, Last sequence update)
DT 01-MAR-2004 (TREMBLER, 26, Last annotation update)
DE Non-phosphorylating glycerol dehydrogenase.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A. PubMed=1854208;
RX MEDLINE=21843113; PubMed=1854208;
RX DOI=10.1128/IAI.70.3.1254-1259.2002;
RA Hughes M.J., Moore J.C., Lane J.D., Wilson R., Pribul P.K.,
RA Younes Z.N., Dobson R.J., Everest P., Reason A.J., Redfern J.M.,
RA Greer F.M., Paxton T., Panico M., Morris H.R., Feldman R.G.,

Q8L5J9 streptococ
Q7NBX5 streptococ
Q9PQC9 streptococ
Q7A5E6 streptococ
Q75TD2 streptococ
Q8ZVJ1 streptococ
DHAL_BACST
YELL_METUA
Q97XAS streptococ
Q976X5 streptococ
Q8EMK2 streptococ
Q97X89 streptococ
Q8CV96 streptococ
Q27059 streptococ
Q8NMB0 streptococ
Q9HUR4 streptococ
Q6M2H6 streptococ
Q65NXX streptococ
Q8U2S5 streptococ
Q8E573 streptococ
Q59702 streptococ
Q81QX6 streptococ
Q6HJ76 streptococ
Q738S1 streptococ
Q9KC36 streptococ
Q8TWC7 streptococ
Q63BQ7 streptococ
P25526 escherichia
Q721Q0 streptococ
Q8RFB1 streptococ
Q73EK2 streptococ
Q81ZE2 streptococ
Q6HP75 streptococ
P42329 streptococ
Q92DA3 streptococ
Q63GQ3 streptococ
Q8Y819 streptococ
Q8X950 streptococ
Q9HK01 streptococ
Q81DV8 streptococ
Q9H0Z2 streptococ
Q57693 streptococ
Q8EMH9 streptococ
Q81IM7 streptococ
Q8BLI8 streptococ
Q9HMT6 streptococ
Q8XOP9 streptococ
Q7VSP3 streptococ
Q7W373 streptococ

RA Santangelo J.D.;
 RT "Identification of major outer surface proteins of *Streptococcus*
 agalactiae.";
 RL Infect. Immun. 70:1254-1259(2002).
 DR EMBL: AF439646; AAL85685.1; -.
 DR HSSP: O59931; 1EWH.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyd_dehydros.
 DR Pfam: PF00171; Aldehyd. 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 SQ SEQUENCE 475 AA; 51198 MW; C525151C88B6AC01 CRC64;

Query Match 99.9%; Score 2391; DB 2; Length 475;
 Best Local Similarity 99.8%; Pred. No. 5,5e-135;
 Matches 474; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTKETONTYNGEWMKSSVNOIEILSPIDSSIGFVPAMTREBYDHAKAGREALPMAALT 60
 :|||||
 DB 1 MTKETONTYNGEWMKSSVNOIEILSPIDSSIGFVPAMTREBYDHAKAGREALPMAALT 60
 QY 61 YVERAOYLHKADIIERDKETATVLAKEISKAYNASVTEVTRADLIRYAABEGIRLT 120
 :|||||
 DB 61 YVERAOYLHKADIIERDKETATVLAKEISKAYNASVTEVTRADLIRYAABEGIRLT 120
 QY 121 SADEGKMDASTGHKLAVIRROPVGIIVLAIPYNPVNLSSGIAPALIGNVVMFKPPT 180
 :|||||
 DB 121 SADEGKMDASTGHKLAVIRROPVGIIVLAIPYNPVNLSSGIAPALIGNVVMFKPPT 180
 QY 122 SADEGKMDASTGHKLAVIRROPVGIIVLAIPYNPVNLSSGIAPALIGNVVMFKPPT 180
 :|||||
 DB 122 SADEGKMDASTGHKLAVIRROPVGIIVLAIPYNPVNLSSGIAPALIGNVVMFKPPT 180
 QY 181 QGSVSGVLAKAPAEAGLPAGVFNITIGRGEISGDIYVEHEEVNFINTGSTPVGQRIK 240
 :|||||
 DB 181 QGSVSGVLAKAPAEAGLPAGVFNITIGRGEISGDIYVEHEEVNFINTGSTPVGQRIK 240
 QY 241 LAGMPPIMLGKDGAGIIVLADADLNAAKOIVAGAYDSGORCTAIRVIVVEEVADEL 300
 :|||||
 DB 241 LAGMPPIMLGKDGAGIIVLADADLNAAKOIVAGAYDSGORCTAIRVIVVEEVADEL 300
 QY 301 AEKISENVAKLSVGPDPFNATVTPVIDNSADFIESTLVVDAROKAKELNEFKDGRILT 360
 :|||||
 DB 301 AEKISENVAKLSVGPDPFNATVTPVIDNSADFIESTLVVDAROKAKELNEFKDGRILT 360
 QY 361 PELFDHVTLDKLAWEPPFGPILPIIRVKDAEBAVAIANKSDPGLQSSVFTRDFOKAFDI 420
 :|||||
 DB 361 PELFDHVTLDKLAWEPPFGPILPIIRVKDAEBAVAIANKSDPGLQSSVFTRDFOKAFDI 420
 QY 421 ANKLEVGTVHINNKTGRGPDNPFPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475
 :|||||
 DB 421 ANKLEVGTVHINNKTGRGPDNPFPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475

RESULT 2

Q8E0B1 PRELIMINARY; PRT; 475 AA.
 AC Q8E0B1;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 GN Glyceraldehyde-3-phosphate dehydrogenase, NADP-dependent.
 OS Name=gagN; OrderedlocusNames=SA080823;
 OS Streptococcus agalactiae (serotype V).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCB1_TaxID=216466;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2603 V/R / Serotype V;
 MEDLINE=22222388; PubMed=12200547; DOI=10.1073/pnae.182380799;
 RA Tetselin H., Masiagnan V., Cieleszewska M.J., Eisen J.A., Peterson S.N.,
 RA Wessels M.R., Paulsen I.T., Nelson K.B., Margalit I., Read T.D.,
 RA Madoff L.C., Wolf A.M., Beaman M.J., Brinkac L.M., Daugherty S.C.,
 RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,

RA Carly H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
 RA Jacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
 RA Rinaldo D., Rappapoli R., Telford J.L., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 emerging human pathogen, serotype V *Streptococcus agalactiae*,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
 DR EMBL: AB014229; AAM99710.1; -.
 DR HSSP: O59931; 1EWH.
 DR TIGR: SA080823; -.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyd_dehydros.
 DR Pfam: PF00171; Aldehyd. 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 475 AA; 51218 MW; F91C086696BAE668 CRC64;

Query Match 99.5%; Score 2382; DB 2; Length 475;
 Best Local Similarity 99.2%; Pred. No. 1.9e-134;
 Matches 471; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LTKETONTYNGEWMKSSVNOIEILSPIDSSIGFVPAMTREBYDHAKAGREALPMAALT 60
 :|||||
 DB 1 MTKETONTYNGEWMKSSVNOIEILSPIDSSIGFVPAMTREBYDHAKAGREALPMAALT 60
 QY 61 YVERAOYLHKADIIERDKETATVLAKEISKAYNASVTEVTRADLIRYAABEGIRLT 120
 :|||||
 DB 61 YVERAOYLHKADIIERDKETATVLAKEISKAYNASVTEVTRADLIRYAABEGIRLT 120
 QY 121 SADEGKMDASTGHKLAVIRROPVGIIVLAIPYNPVNLSSGIAPALIGNVVMFKPPT 180
 :|||||
 DB 121 SADEGKMDASTGHKLAVIRROPVGIIVLAIPYNPVNLSSGIAPALIGNVVMFKPPT 180
 QY 122 SADEGKMDASTGHKLAVIRROPVGIIVLAIPYNPVNLSSGIAPALIGNVVMFKPPT 180
 :|||||
 DB 122 SADEGKMDASTGHKLAVIRROPVGIIVLAIPYNPVNLSSGIAPALIGNVVMFKPPT 180
 QY 181 QGSVSGVLAKAPAEAGLPAGVFNITIGRGEISGDIYVEHEEVNFINTGSTPVGQRIK 240
 :|||||
 DB 181 QGSVSGVLAKAPAEAGLPAGVFNITIGRGEISGDIYVEHEEVNFINTGSTPVGQRIK 240
 QY 241 LAGMPPIMLGKDGAGIIVLADADLNAAKOIVAGAYDSGORCTAIRVIVVEEVADEL 300
 :|||||
 DB 241 LAGMPPIMLGKDGAGIIVLADADLNAAKOIVAGAYDSGORCTAIRVIVVEEVADEL 300
 QY 301 AEKISENVAKLSVGPDPFNATVTPVIDNSADFIESTLVVDAROKAKELNEFKDGRILT 360
 :|||||
 DB 301 AEKISENVAKLSVGPDPFNATVTPVIDNSADFIESTLVVDAROKAKELNEFKDGRILT 360
 QY 361 PELFDHVTLDKLAWEPPFGPILPIIRVKDAEBAVAIANKSDPGLQSSVFTRDFOKAFDI 420
 :|||||
 DB 361 PELFDHVTLDKLAWEPPFGPILPIIRVKDAEBAVAIANKSDPGLQSSVFTRDFOKAFDI 420
 QY 421 ANKLEVGTVHINNKTGRGPDNPFPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475
 :|||||
 DB 421 ANKLEVGTVHINNKTGRGPDNPFPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475

RESULT 3

Q8E5Y8 PRELIMINARY; PRT; 475 AA.
 AC Q8E5Y8;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 GN Hypothetical protein gbs0841.
 OS OrderedlocusNames=gbs0841;
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCB1_TaxID=216495;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEM316 / Serotype III;
 MEDLINE=22242508; PubMed=12354221;

RA Glaeser P., Rusniok C., Buchrieser C., Chevallier F., Frangoul L.,
RA Maeder T., Zouine M., Couve E., Lailou L., Poyart C., Trieru-Quot P.,
RA Kunst F.,
RT "genome sequence of *Streptococcus agalactiae*, a pathogen causing
RT invasive neonatal disease."
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL: AL766847; CAD46485.1; -
DR HSSP: Q59931; 1EUN.
DR Sagalier: sbe0841; -
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002086; Aldehyd_dehydrog.
DR Pfam: PF00171; Aldehyd. 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR Complete proteome.
SQ SEQUENCE 475 AA; 51217 MW; DE33AE79B81A66B6 CRC64;

Query Match 99.4%; Score 2379; DB 2; Length 475;
Best Local Similarity 99.2%; Pred. No. 2.9e-134;
Matches 471; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LTKYQNYVNGEKSSVNOIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAAALYT 60
Db 1 MTKYQNYVNGEKSSVNOIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAAALYT 60

QY 61 VYRAQYLHKADIIERDKKEIATVLAKEISKAYNASVTEVVTADLIRYAAEGIRLST 120
Db 61 VYRAQYLHKADIIERDKKEIATVLAKEISKAYNASVTEVVTADLIRYAAEGIRLST 120

QY 121 SABEGGMASSTGHKAVIRPROVGIYLAIPNYNPVNLGSKIAPALIGGNVMPKPP 180
Db 121 SABEGGMASSTGHKAVIRPROVGIYLAIPNYNPVNLGSKIAPALIGGNVMPKPP 180

QY 181 QGSVSGVLVAKAFEAAGLPAGVFNITGRGSEIGDYVEHEEVNFINTGSTPVGORIG 240
Db 181 QGSVSGVLVAKAFEAAGLPAGVFNITGRGSEIGDYVEHEEVNFINTGSTPVGORIG 240

QY 241 LAQMRPIMLELGKDGAGIVLADADLNNAKQIVAGAYDSGQCTAIKRYLVVEEVADEL 300
Db 241 LAQMRPIMLELGKDGAGIVLADADLNNAKQIVAGAYDSGQCTAIKRYLVVEEVADEL 300

QY 301 AEKISENVAKLSVGDPDNATVPTVIDNSADFIESIYVDARQGAKEINEFRDGRILT 360
Db 301 AEKISENVAKLSVGDPDNATVPTVIDNSADFIESIYVDARQGAKEINEFRDGRILT 360

QY 361 PGHFDHTVLDKLAWEPEFGPIIPRIKVDABEAVAIANKSPGLQSSVFTTRDFOKAFDI 420
Db 361 PGHFDHTVLDKLAWEPEFGPIIPRIKVDABEAVAIANKSPGLQSSVFTTRDFOKAFDI 420

QY 421 ANKLEVGTVHINKKTGRGPDNPFPLGLKSGAGVQIGIRYSIEAMTVKSIIVLDMK 475
Db 421 ANKLEVGTVHINKKTGRGPDNPFPLGLKSGAGVQIGIRYSIEAMTVKSIIVLDMK 475

RESULT 4
Q97QTO :
ID Q97QTO PRELIMINARY; PRT; 474 AA.
AC Q97QTO;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase, NADP-dependent.
GN OrderedLocustNames=SP119;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
ON NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; Pubmed=11463916; DOI=10.1126/science.1061217;
Tetelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,

RA Peterson S.N., Heidelberg J.F., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E.K., Knouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of *Streptococcus*
RT pneumoniae."
RL Science 293:498-506 (2001).
DR EMBL: AE007413; AAK75230.1; -
DR PIR: E95129; E95129.
DR HSSP: Q59931; 1EUN.
DR TIGR: SP1119; -
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002086; Aldehyd_dehydrog.
DR Pfam: PF00171; Aldehyd. 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR Complete proteome.
SQ SEQUENCE 474 AA; 51127 MW; 89D8C05F02AC0048 CRC64;

Query Match 75.7%; Score 1812; DB 2; Length 474;
Best Local Similarity 74.1%; Pred. No. 2.4e-100;
Matches 349; Conservative 57; Mismatches 65; Indels 0; Gaps 0;

QY 5 YQNYVNGEKSSVNOIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAAALTYVER 64
Db 4 YQNYVNGEKSSVNOIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAAALTYVER 64

QY 65 AOTLHKADIIERDKKEIATVLAKEISKAYNASVTEVVTADLIRYAAEGIRLSTADE 124
Db 65 AOTLHKADIIERDKKEIATVLAKEISKAYNASVTEVVTADLIRYAAEGIRLSTADE 124

QY 125 GGMASSTGHKAVIRPROVGIYLAIPNYNPVNLGSKIAPALIGGNVMPKPP 184
Db 125 GGMASSTGHKAVIRPROVGIYLAIPNYNPVNLGSKIAPALIGGNVMPKPP 184

QY 184 GGFETSKKKAIVAREBVGIVLAPNYNPVNLGSKIAPALIGGNVMPKPP 183
Db 184 GGFETSKKKAIVAREBVGIVLAPNYNPVNLGSKIAPALIGGNVMPKPP 183

QY 245 RPIIMLELGKDGAGIVLADADLNNAKQIVAGAYDSGQCTAIKRYLVVEEVADELAEKI 304
Db 245 RPIIMLELGKDGAGIVLADADLNNAKQIVAGAYDSGQCTAIKRYLVVEEVADELAEKI 304

QY 305 SENVAKLSVGDPDNATVPTVIDNSADFIESIYVDARQGAKEINEFRDGRILT 364
Db 305 SENVAKLSVGDPDNATVPTVIDNSADFIESIYVDARQGAKEINEFRDGRILT 364

QY 364 QEVSKLIVGDPDNATVPTVIDNSADFIESIYVDARQGAKEINEFRDGRILT 363
Db 364 QEVSKLIVGDPDNATVPTVIDNSADFIESIYVDARQGAKEINEFRDGRILT 363

QY 425 EVGTVHINKKTGRGPDNPFPLGLKSGAGVQIGIRYSIEAMTVKSIIVLDMK 475
Db 425 EVGTVHINKKTGRGPDNPFPLGLKSGAGVQIGIRYSIEAMTVKSIIVLDMK 475

RESULT 5
Q8DP57 :
ID Q8DP57 PRELIMINARY; PRT; 474 AA.
AC Q8DP57;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC
DE 1.2.1.9).
GN Name=gadN; OrderedLocustNames=SP11028;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

OK NCBI_TaxID=171101;
 RN [1]
 RX SEQUENCE FROM N.A.
 MEDLINE=21429245; PubMed=11544234;
 DOI=10.1128/JB.183.19.5709-5717.2001;
 RA Hoskins J., Alborn W.E. Jr., Arnold U., Blaszcak L.C., Burgett S.,
 Dehoff B.S., Battem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 LeBlanc D.J., Lee L.N., Lefkowitz B.J., Lu J., Matsushita P.,
 McLaren S.M., McHenry M., Mclester K., Mundy C.W., Nicas T.I.,
 Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
 Sun P.-M., Winkler M.B., Yang Y., Young-Bellido M., Zhao G.,
 Zook C.A., Baltz R.H., Jaskunas S.R., Kosteck P.R. Jr., Skatrud P.L.,
 RA Glaes J.I.;
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
 RL J. Bacteriol. 183:5709-5717(2001).
 DR EMBL, AE008475; AKK9832.1; -.
 DR PIR, D98000; D98000.
 DR HSSP; Q5931; 1E0H.
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldehyde_dehydratase.
 DR Pfam; PF00171; Aldehyd_1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW complete proteome.
 SQ SEQUENCE 474 AA; 51076 MW; 87BC0A8BD7E5E21 CRC64;

Query Match 75.3%; Score 1802; DB 2; Length 474;
 Best Local Similarity 73.7%; Pred. No. 9.5e-100;

Matches 347; Conservative 58; Mismatches 66; Indels 0; Gaps 0;

QY 5 YONTYNGEWSVNOQIEILSPIDDSLGFPVPMTRREVDHAKAGREAPAMALTVYR 64
 DB 4 YONLVNGKSKSSBQELTITSPINOELGTVPMTOFDEANQAAALPAMRALSATR 63
 QY 65 AQLYKADIIERDKKEIATVLAKEISKAYNSVTEVETADILIRYAEEGRLSTADE 124
 DB 64 AAVLHKTAAILERDKKEIGITIAKEVAKGKKAIGEVETADILIRYAEEGRLSTADE 123
 QY 125 GGMDSSTGHKLAVIRORVGVLAIPNYPNLSGSKTAPRLIGANNVMPFPQSGV 184
 DB 124 GGFPEASAKKLAVRREBPVGLAIPNYPNLSASKTAPRLAIGNVMMKPPQSGSI 183
 QY 185 SGLVLAFAAEAGLPAGVFNITGRGSEIGDYVEHEEVNFINTGSTPVGRIGLAGM 244
 DB 184 SGLLAKAFEBEAGIPAGVFNITGRGSEIGDYIIEHKEVNFINTGSTPVGRIGLAGM 243
 QY 245 RPTMELGSKDAGIVLADADLNAHQIVAGAYDYSGORCTAIKRVLVVEVADELAEXI 304
 DB 244 RPTMELGSKDAGIVLADADLNAHQIVAGAYDYSGORCTAIKRVLVVEVADELAEXI 303
 QY 305 SERVVALSTGDDPPDNATVTPVINDNADPTESLVVARQKAKELNEFKDGLLTGIF 364
 DB 304 QEVSVALSTGDDPPDNADITPVINDNADPTESLVVARQKAKELNEFKDGLLTGIF 363
 QY 365 DHVTLDMKLAMEEPPFPILPIIRVKDAEVAIVANKSDGLGSSVTRPDPFOKAFDIANL 424
 DB 364 DQVTKMKVAMEEPPFPVLPPIIRVAVVEEIAIAPANSSEGLGSSVTRPDPFOKAFDIANL 423
 QY 425 EVGTVAHINKTGKGPDPNPFPLGLKSGAGVQIRYSIEANTVKSIVLDMK 475
 DB 424 EVGTVAHINKTGKGPDPNPFPLGLKSGAGVQIRYSIEANTVKSIVLDMK 474

RESULT 6
 ID GABN_STRMU STANDARD; PRT; 475 AA.
 AC 059931;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9)

DE (Non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase)
 DE (glyceraldehyde-3-phosphate dehydrogenase [NADP+]) (triosephosphate
 dehydrogenase).
 GN Name=gapn; OrderedLocustNames=SMU.676;
 OS Streptococcus mutans; Lactobacillales; Streptococcaceae;
 OC Bacteria; Firmicutes; Streptococcus;
 OC Streptococcus.
 NC NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NGS / Serotype C;
 RA MEDLINE=95270576; PubMed=7751269;
 RX Boyd D.A., Cvitkovitch D.G., Hamilton I.R.;
 RT "Sequence, expression, and function of the gene for the
 RT nonphosphorylating, NADP-dependent glyceraldehyde-3-phosphate
 RT dehydrogenase of Streptococcus mutans.";
 RL J. Bacteriol. 177:2622-2627(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 CC -I- FUNCTION: Important as a means of generating NADPH for
 CC biosynthetic reactions.
 CC -I- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + NADP(+) + H(2)O
 CC = 3-phospho-D-glycerate + NADPH.
 CC -I- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC
 DR EMBL, L38521; AAA91091.1; -.
 DR EMBL, AB014911; AAN58410.1; -.
 DR PIR; A57151; A57151.
 DR PDB; 1E0H; X-ray; A/B/C/D=1-475.
 DR PDB; 1Q11; X-ray; A/B/C/D=1-475.
 DR PDB; 1Q16; X-ray; A/B/C/D=1-475.
 DR PDB; 2E0H; X-ray; A/B/C/D=1-475.
 DR InterPro; IPR002086; Aldehyde_dehydratase.
 DR Pfam; PF00171; Aldehyd_1.
 DR TIGRfam; TIGR01804; BADH; 1.
 DR TIGRfam; TIGR01722; MMSDH; 1.
 DR TIGRfam; TIGR01780; SSADH; 1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; FALSE NEG.
 KW 3D-structure: Complete proteome: NADP; Oxidoreductase.
 KW NP BIND 230 235
 FT ACT_SITE 250 250
 FT ACT_SITE 284 284
 FT CONFLICT 58 58
 FT CONFLICT 85 85
 FT CONFLICT 347 347
 FT STRAND 4 4
 FT STRAND 7 9
 FT TURN 10 11
 FT STRAND 12 14
 FT STRAND 19 23
 FT TURN 25 27
 FT STRAND 30 35
 FT HELIX 39 58
 FT HELIX 61 77
 FT TURN 78 78

FT	HELIX	79	90
FT	TURN	91	91
FT	HELIX	94	115
FT	TURN	116	117
FT	STRAND	120	124
FT	HELIX	125	127
FT	TURN	128	128
FT	HELIX	130	132
FT	TURN	133	134
FT	STRAND	135	143
FT	STRAND	146	150
FT	TURN	153	154
FT	TURN	156	157
FT	HELIX	158	169
FT	TURN	170	171
FT	STRAND	173	177
FT	TURN	181	182
FT	HELIX	183	196
FT	TURN	197	197
FT	TURN	200	202
FT	STRAND	203	205
FT	HELIX	210	219
FT	TURN	221	222
FT	STRAND	225	229
FT	HELIX	232	241
FT	TURN	242	244
FT	STRAND	247	250
FT	STRAND	255	259
FT	TURN	261	262
FT	HELIX	265	277
FT	HELIX	278	281
FT	STRAND	287	293
FT	HELIX	294	309
FT	TURN	310	310
FT	STRAND	313	313
FT	HELIX	316	318
FT	TURN	319	319
FT	STRAND	322	322
FT	HELIX	328	343
FT	TURN	344	345
FT	STRAND	347	348
FT	STRAND	354	355
FT	TURN	356	357
FT	STRAND	358	359
FT	STRAND	362	365
FT	TURN	369	370
FT	HELIX	372	374
FT	TURN	375	375
FT	STRAND	382	388
FT	HELIX	391	400
FT	STRAND	403	404
FT	STRAND	406	410
FT	HELIX	414	423
FT	STRAND	428	431
FT	TURN	440	441
FT	STRAND	444	444
FT	STRAND	447	448
FT	TURN	449	450
FT	STRAND	451	452
FT	STRAND	455	455
FT	HELIX	457	463
FT	TURN	464	464
FT	STRAND	465	473
SO	SEQUENCE	475 AA;	51194 MW; FOA2770AB9552DC CRC64;

Query Match 71.0%; Score 1700; DB 1; Length 475;
Best Local Similarity 66.9%; Pred. No. 1.2e-93;
Matches 318; Conservative 75; Mismatches 82; Indels 0; Gaps 0;

Qy 1 LTRKYQYVNGEMKSSVNOIEILSPIDSSLGFPVPMTRREVDHAMKAGREALPMAAALT 60
Db 1 MTKQYKNLVNGEMKLSENEIKIYEPASGALGVPAMSTEEVDYVYASAKKQAPAKRSLS 60

Qy	61	YVERAOYLHKADIIERDKKEIATVLAKEISKAYNASVTEVVRTADLIRYAAEGIRLST 120
Db	61	YIERAAVYIHKVADIIIRDKKEIGAVISKVANGYSAVSEVVRTAIIYVAAEGIRMEG 120
Qy	121	SADEGKMDASTGSHKLAVIRQPVGIVLAIAEPYVFNISGSKIAPALIGVYVWPKPPT 180
Db	121	EVLGEGSFPAASKKIAVRRREPVGIVLAISFPNPNVLAGSKIAPALIAGVIAKPPPT 180
Qy	181	QGSVSGVLAKAPAEAGDPAGVFNITITGSGSISGDIYVEHEVNFINTGSPVQGRICK 240
Db	181	QGISGGLIAEAPAEAGLPAGVFNITITGSGSEIGDIYVHQVNFINTGSGIGIRICK 240
Qy	241	LAGMRPIMLEIGKQAGITLADADLDMAKQIYAGAYDVSQGRCTAIKRVVVEEYADL 300
Db	241	MAGMRPIMLEIGKQSAIYLEDADLETAKNIIAGFGYSGQRTAVKVLVWESYADL 300
Qy	301	AEKISENVAKLSGDPFDNAVATVPVDDVSADPFIESLVYDADOKAKEINPEADGRLLT 360
Db	301	VEKIREKVLALITIGNPEDDADITPLIDTSADYVEGLINDADKGAALTEIKREGNLIC 360
Qy	361	PGLFPHVTLDMLKAMEEPPGPILPIIRYDAEBAVAIAANKSDPGLQSVFTSDPQAPDI 420
Db	361	PIIFDKVTTDMRLAMEEPPGPVLPPIIRYTSVEEALIESNKSEYGLQASIFNDPFPAPGI 420
Qy	421	ANKLEVGTVHINKTGRGPDNPFFLGKSGAGVQGIRISITAMTVKSIYVDMK 475
Db	421	ASQLEVGTVHINKTKRGTDNPPFLGAKKSGAGIQGVKYSIEAMTVKSVVDIK 475

RESULT 7
Q8POF5 PRELIMINARY; PRT; 475 AA.
ID Q8POF5;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Putative NADP-dependent glyceroldehyde-3-phosphate dehydrogenase.
GN Name=gadN; Ordered locus names=spym18_1983;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232;
RA MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.B., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AE010058; AAL97978.1; -.
DR HSSP; Q59931; 1E0H.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde_dehydrog.
DR Pfam; PF00171; Aldehdh_1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR Complete proteome.
SO SEQUENCE 475 AA; 50381 MW; 335E29B77212DFC5 CRC64;

Query Match 68.7%; Score 1645; DB 2; Length 475;
Best Local Similarity 65.9%; Pred. No. 2.4e-90;
Matches 313; Conservative 75; Mismatches 87; Indels 0; Gaps 0;

Qy 1 LTRKYQYVNGEMKSSVNOIEILSPIDSSLGFPVPMTRREVDHAMKAGREALPMAAALT 60
Db 1 MAKQYKNLVNGEMKLSENEITVYAPATGEBELGVPAMTQAEVDAYASAKKALSDBWRAIS 60
Qy 61 YVERAOYLHKADIIERDKKEIATVLAKEISKAYNASVTEVVRTADLIRYAAEGIRLST 120

```
Db      61 YVERAAVYHKADILVRDAEKIGALISKEVAKGHKAASVEVIRTAETIINYAAEEGRLMEG 120
      121 SADEGKMDSYTGHTKLAIVRRQVGI VLAIAPYNYPVNI.SGSKIAFALIGANVVMFKPPT 180
      122 EVLGGSGFEASAKKKIAIVRRBPVGLVLAISPNNYPVNLASKIAFALAGNVVALKRPPT 180
      121 EVLGGSGFEASAKKKIAIVRRBPVGLVLAISPNNYPVNLASKIAFALAGNVVALKRPPT 180
      181 QGSVSGVLAKAPAEAGLPAGVNTTIGRSGEIGDYIVHEEYVNFINTGSTPVGORIGK 240
      181 QGSISGLLAEAFABAGIPAGVNTTIGRSGVIGDYIVHEEYVNFINTGSTPVGORIGK 240
      241 LAGMRPIMELGKGKSAIVLEBDADLAAKNIVAGFVSGQRCYAKVVLVMDKYADQL 300
      241 LAGMRPIMELGKGKSAIVLEBDADLAAKNIVAGFVSGQRCYAKVVLVMDKYADQL 300
      301 AEKISENVAKLSTGVPDFNATVTPVYIDNSADFIETSLVVDARQKAKELNEFKDGRLLT 360
      301 AAEIKTLVKLSVGNPEDDADITPLIDTSADFPVEGLIKDADTKGATLTAFFNRGNLIS 360
      361 PGLFDHVTLDMLKLAWEPPGPPLPIIRVKDAEBAVAIANKSPDGLQSSVFTTRDFOKAFDI 420
      361 PVLFDHVTLDMLKLAWEPPGPPLPIIRVTVVEBAVKISNESEYGLQASIFFTNFPKAFGI 420
      421 ANKLEVTGTHINNKTRGPDNPPFLGLKSGAGVQGIKYSIEAMTVKSIYVDMK 475
      421 AEOLSVGVTHLNKTKRGTDNFPFLGAKKSGAGVQGVKYSIEAMTVKSVVFDIQ 475
```

RESULT 8

PRELIMINARY, PRT, 475 AA.

```
AC 099267; PRELIMINARY, PRT, 475 AA.
AC 099267;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9).
GN Name=gapN; OrderedlocusNames=SPY1371;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RC MEDLINE=21192664; PubMed=1236296; DOI=10.1073/pnas.151559398;
RA Perrecci J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Szecze S., Suvorov A.N., Kenon S., Lai H.S., Lin S.P.,
RA Qian Y., Clifton S.W., Roe B.A., McLaughlin R.E.;
RA "Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;"
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.;"
RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
RL EMBL: AB006575; AAK34198.1; -.
DR HSPF; Q59931; IEUH.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 475 AA; 50368 MW; BC24957621AE0FC9 CRC64;
```

Query Match 68.5%; Score 1639; DB 2; Length 475;

Best Local Similarity 65.5%; Pred. No. 5.4e-90;

Matches 311; Conservative 77; Mismatches 87; Indels 0; Gaps 0;

```
QY      1 LTKEXONYNGEKSSVNOIEILSPIDSSLGFPVPMTRBEVDHANKAGREALPAAALYT 60
      1 LTKEXONYNGEKSSVNOIEILSPIDSSLGFPVPMTRBEVDHANKAGREALPAAALYT 60
      1 MAKQYKLVNGEKLSENIEITTYAATGSELGVPAMTQAEVDVAVASAKKALSDWRALS 60
      1 MAKQYKLVNGEKLSENIEITTYAATGSELGVPAMTQAEVDVAVASAKKALSDWRALS 60
      61 VVERAOYLHKADIIIRDKSEIATVLAKEISKAYNASVTEVATADLIRYAAEEGRLST 120
      61 VVERAOYLHKADIIIRDKSEIATVLAKEISKAYNASVTEVATADLIRYAAEEGRLST 120
      61 YVERAAVYHKADILVRDAEKIGALISKEVAKGHKAASVEVIRTAETIINYAAEEGRLMEG 120
      61 YVERAAVYHKADILVRDAEKIGALISKEVAKGHKAASVEVIRTAETIINYAAEEGRLMEG 120
```

```
QY      121 SADEGKMDSYTGHTKLAIVRRQVGI VLAIAPYNYPVNI.SGSKIAFALIGANVVMFKPPT 180
      122 EVLGGSGFEASAKKKIAIVRRBPVGLVLAISPNNYPVNLASKIAFALAGNVVALKRPPT 180
      121 EVLGGSGFEASAKKKIAIVRRBPVGLVLAISPNNYPVNLASKIAFALAGNVVALKRPPT 180
      181 QGSVSGVLAKAPAEAGLPAGVNTTIGRSGEIGDYIVHEEYVNFINTGSTPVGORIGK 240
      181 QGSISGLLAEAFABAGIPAGVNTTIGRSGVIGDYIVHEEYVNFINTGSTPVGORIGK 240
      241 LAGMRPIMELGKGKSAIVLEBDADLAAKNIVAGFVSGQRCYAKVVLVMDKYADQL 300
      241 LAGMRPIMELGKGKSAIVLEBDADLAAKNIVAGFVSGQRCYAKVVLVMDKYADQL 300
      301 AEKISENVAKLSTGVPDFNATVTPVYIDNSADFIETSLVVDARQKAKELNEFKDGRLLT 360
      301 AAEIKTLVKLSVGNPEDDADITPLIDTSADFPVEGLIKDADTKGATLTAFFNRGNLIS 360
      361 PGLFDHVTLDMLKLAWEPPGPPLPIIRVKDAEBAVAIANKSPDGLQSSVFTTRDFOKAFDI 420
      361 PVLFDHVTLDMLKLAWEPPGPPLPIIRVTVVEBAVKISNESEYGLQASIFFTNFPKAFGI 420
      421 ANKLEVTGTHINNKTRGPDNPPFLGLKSGAGVQGIKYSIEAMTVKSIYVDMK 475
      421 AEOLSVGVTHLNKTKRGTDNFPFLGAKKSGAGVQGVKYSIEAMTVKSVVFDIQ 475
```

RESULT 9

PRELIMINARY, PRT, 467 AA.

```
AC 08K707; PRELIMINARY, PRT, 467 AA.
AC 08K707;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase.
GN Name=gapN; OrderedlocusNames=SPYM3_1045;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RC MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schliwert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.;"
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
RL EMBL: AB014157; AAM79652.1; -.
DR HSPF; Q59931; IEUH.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 467 AA; 49478 MW; 287AEDC6963B82EB CRC64;
```

Query Match 67.6%; Score 1618; DB 2; Length 467;

Best Local Similarity 66.0%; Pred. No. 9.5e-89;

Matches 308; Conservative 74; Mismatches 85; Indels 0; Gaps 0;

```
QY      9 VNGEKSSVNOIEILSPIDSSLGFPVPMTRBEVDHANKAGREALPAAALTVVERAOYL 68
      9 VNGEKSSVNOIEILSPIDSSLGFPVPMTRBEVDHANKAGREALPAAALTVVERAOYL 68
      1 NNGEKLSENIEITTYAATGSELGVPAMTQAEVDVAVASAKKALSDWRALSYVERAAYL 60
      1 NNGEKLSENIEITTYAATGSELGVPAMTQAEVDVAVASAKKALSDWRALSYVERAAYL 60
      69 HKAADIIIRDKSEIATVLAKEISKAYNASVTEVATADLIRYAAEEGRLSTSADEGKGM 128
      69 HKAADIIIRDKSEIATVLAKEISKAYNASVTEVATADLIRYAAEEGRLSTSADEGKGM 128
      61 HKAADILVRDAEKIGALISKEVAKGHKAASVEVIRTAETIINYAAEEGRLMEGEVLGGSGF 120
      61 HKAADILVRDAEKIGALISKEVAKGHKAASVEVIRTAETIINYAAEEGRLMEGEVLGGSGF 120
```

[illegible]

QY	123	BEGGMDASTGHKLAVIRROPVGIATAPPNYPNLSGGSKAPLALIGNNVMFKPPTQG	182
D5	131	INSDNPFGSKDKDCKSLVERVPGLVIATLSPPNYPNLSSSKAAPALIAAGNSVLKPKSTGT	190
QY	183	SVSGIVLKAAFAEAGLPAGVFNTTIGRSGSEIDGYIVEHEBNVFIIFTGSTPYGORIGKLA	242
D6	191	AISALHLIAETINNAAGLPAGVANTVTYGKSEICGDYILTHEEVNFIFNTGSSANGKHISKIA	250
QY	243	GMRPIIMTELGGKDAGITVLADDLDNAKOIYAGAAYDSGQRCTAIKRVLVEEVADELAE	302
D6	251	GMIPIVTELGKGDAIVLEADANLETAKSIVSAGAYSGQRCTAIVKRVLMVKVDELYE	310
QY	303	KISENVALTASGDPEDNATMTPPVVIDDNDAFIESLVADAOKGAELNEFFKDDGLLTG	362
D6	311	LVTKKVKSELKGNPFDDVTTIPPLIDNKADVVQTLLIDDAIEKGATLIVNKKRENIMLYPT	370
QY	363	LFDHVTLDMKLAMEEPFPILPIIRVKCAEEVAIALANKSPDFLOSGSVTFPRDFOKAFDIAN	422
D6	371	LFENVTLADMRLAMEEPFGPVLPPIRKVSMDAIELANSSEYLOQSAVFTENMHDAFYIAN	430
QY	423	KLEVGVTHINKKTGRGPDNPFLPLGKSQAGVGQIRYSISTEAMTNVKSIVLDM	474
D6	431	KLDVGTVOVNKNPERGPDHPFLPGTKSSGMGTGQIRYSISTEAMTRHKSIVLNL	482
RESULT 11			
Q73CX4		PRELIMINARY; PRT; 479 AA.	
ID AC Q73CX4;			
DT 05-JUL-2004 (TREMBLrel. 27, Created)			
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)			
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)			
DE Glyceraldehyde-3-phosphate dehydrogenase, NADP-dependent (EC 1.2.1.9)			
GN Name=gapN; OrderedLocustNames=CE00940;			
OS Bacillus cereus (strain ATCC 10987).			
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX NCBI_taxonomy=222523;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX PubMed=14960714; DOI=10.1093/nar/gkh258;			
RA Raeko D.A., Ravel J., Oekstrad O.A., Helgason E., Cer R.Z., Jiang L., Shores K.A., Fouts D.E., Tourasse N.J., Angilou S.V., Kolonay J.F., Nelson W.C., Kolstoe A.B., Fraser C.M., Read T.D.;			
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic adaptations and a large plasmid related to Bacillus anthracis pXOI."			
RL Nucleic Acids Res. 32:977-988(2004).			
DR EMBL: AB017267; AAC939871.1; -.			
DR HSSP; P56533; IAMS.			
DR TIGER; BCE0940; -.			
DR GO; GO:0016491; F:oxidoreductase activity; IEA.			
DR GO; GO:0008152; P:metabolism; IEA.			
DR InterPro; IPRO02086; Aldehyd_dehydrosug.			
DR Pfam; PF00171; Aldehd; 1.			
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.			
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.			
KW Complete proteome.			
SC SEQUENCE 479 AA; 51798 MW; E294F44609BCA580 CRC64;			
Query Match			
Best Local Similarity	56.2%; Score 1345; DB 2; Length 479;		
Matches	266; Conservative 77; Mismatches 127; Indels 2; Gaps 1		
QY	5 YONTYNGEMK--SSVNOJIEILSPIDDSSLGFVPANTREBEVDHAMKAREALPAMAALTVY	62	
D6	7 YKFYINGEMRESSCGERTIEIPSPYLHEVIGVOAITRGGEVDEAIASAKAOKSMASDAQ	66	
QY	63 ERAOYIHKAADIIRERKEEIATVLAKESIKAYMASVTEVVRADLIIRYAEEGISLSTSA	122	
D6	67 DRARYLYKMADELVMNMDEIIDIMKEVKGKQKARKKVVRADDIRTYITEALHHGES	126	
QY	123 DEGGMDASTGHKLAVIRROPVGIATAPPNYPNLSGGSKAPLALIGNNVMFKPPTQG	182	

```

Db      127 MMGDSFPGGTGSKLAIIGRAPLGVLAIPFNNPVNLSAKLAPALIMGNNAVIFKPAATG 186
Qy      183 SVSGVLAKAFABAGLPAGVFNITTCRGSEIGDYIVHEEENVPINFSTGTPVQGRIGKLA 242
Db      187 AISGIKMVEALHKAQIPKGLVNVATGRSGVIGDYIVHEEGINMVSFTGTTNGKILAKKA 246
Qy      243 GMRPIMLELGGKADGIVLADADLDNNAKOIYAGAYDSGRCCTAIKRYLVVEEVADELAE 302
Db      247 AMIPLVLELGGKDPGIVREDADLDQDAAHHIVSGAFSYSGRCCTAIKRYLVHEVNADELVG 306
Qy      303 KISENVAKISVGDPEPDNATVTPVIDNNSADFIESTLVNAROKAKELNEFPKDGRLITPG 362
Db      307 LIQEQVAKISVSGSPQDSTIVPLIDDKSADFVQGLVDAVEKAGATVIGNKRENNLIYPT 366
Qy      363 LEDHTLDMKLAEEFPFPIIPIIRYKDAEBAVAIANKSDPGLQSSVFTTRDPQKAFDIAN 422
Db      367 LIDHTYEDMKVAMEEFPFPIIPIIRVSSDEQAIETANKSEFGLQASVFTKDIINKAFALAN 426
Qy      423 KLEVGTVHINNKTGRGPNPFPLIGKSGAGVQGRISYEIEMTNTKSIYVLM 474
Db      427 KIETGVOINGRTERGPDHPFPIIGVKGSGMGAQGRKSLSEMTREKVTYVNL 478

```

```

RESULT 12
O6HMX0
ID AC 06HMX0: PRELIMINARY; PRT; 479 AA.
AC 06HMX0:
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (NADP+) (EC 1.2.1.9).
GN Name=gapN; OrderedLocusNames=Bt9727.0757;
OS Bacillus thuringiensis (subsp. konkukian).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxId=180856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-27;
RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Rice H.;
RT "Complete genome sequence of Bacillus thuringiensis 97-27."
RT Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017355; AAT59146.1; -.
DR GO; GO:0008866; F:glyceraldehyde-3-phosphate dehydrogenase (N. .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrotrog.
DR Pfam; PF00171; Aldehyd. 1.
DR PROSITE; PS000670; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.
KW Complete proteome.
SQ SEQUENCE 479 AA; 51858 MW; 679787AF11DB86F7 CRC64;

```

Query Match 56.2%; Score 1344; DB 2; Length 479;
 Best Local Similarity 56.1%; Pred No. 2.4e-72;
 Matches 265; Conservative 78; Mismatches 127; Indels 2; Gaps 1;

```

Qy      5 YONYVNGEWK--SSVNOIEILSPIDDSLGFVPAMTREEDVDHAKMGREALPAMALTVY 62
Db      7 YKFLYNGEMWRSSSGETIEIPSPYLHEVYIGVQVATRGSEVDALASAKAEQSMVAESLQ 66
Qy      63 ERAQYLHKAADIIIEPKKEIATVLAKEISKAYNASVTEVTRADLIRYAEEGIRLSTSA 122
Db      67 DRAKLYKMADELVMQDEIADIIMKEVGKGYKAKKEVTRADFIIRYTIIEALHMGES 126
Qy      123 DEGGKMDASTGKLAIVIRROPVGIYLAIPYNPVNLSSGKIAPALIGNVMPKRPFG 182
Db      127 MMGDSFPGGTGSKLAIIGRAPLGVLAIPFNNPVNLSSAKLAPALIMGNNAVIFKPAATG 186
Qy      183 SVSGVLAKAFABAGLPAGVFNITTCRGSEIGDYIVHEEENVPINFSTGTPVQGRIGKLA 242

```

```

Db      187 AISGIKMVEALHKAQIPKGLVNVATGRSGVIGDYIVHEEGINMVSFTGTTNGKILAKKA 246
Qy      243 GMRPIMLELGGKADGIVLADADLDNNAKOIYAGAYDSGRCCTAIKRYLVVEEVADELAE 302
Db      247 SMIPVLVLELGGKDPGIVREDADLDQDAAHHIVSGAFSYSGRCCTAIKRYLVHEVNADELVS 306
Qy      303 KISENVAKISVGDPEPDNATVTPVIDNNSADFIESTLVNAROKAKELNEFPKDGRLITPG 362
Db      307 LYKEQVAKISVSGSPQDSTIVPLIDDKSADFVQGLVDAVEKAGATVIGNKRENNLIYPT 366
Qy      363 LEDHTLDMKLAEEFPFPIIPIIRYKDAEBAVAIANKSDPGLQSSVFTTRDPQKAFDIAN 422
Db      367 LIDHTYEDMKVAMEEFPFPIIPIIRVSSDEQAIETANKSEFGLQASVFTKDIINKAFALAN 426
Qy      423 KLEVGTVHINNKTGRGPNPFPLIGKSGAGVQGRISYEIEMTNTKSIYVLM 474
Db      427 KIETGVOINGRTERGPDHPFPIIGVKGSGMGAQGRKSLSEMTREKVTYVNL 478

```

```

RESULT 13
O63FP8
ID AC O63FP8: PRELIMINARY; PRT; 479 AA.
AC O63FP8:
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (NADP+) (EC 1.2.1.9).
GN Name=gapN; ORFNames=BtZK0750;
OS Bacillus cereus ZK.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxId=288681;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZK;
RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Rice H.;
RT "Complete genome sequence of Bacillus cereus ZK."
RT Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000001; AAU19493.1; -.
KW Oxidoreductase.
SQ SEQUENCE 479 AA; 51828 MW; DA9209D3A44FB04B CRC64;

```

Query Match 56.0%; Score 1341; DB 2; Length 479;
 Best Local Similarity 56.1%; Pred No. 3.7e-72;
 Matches 265; Conservative 78; Mismatches 127; Indels 2; Gaps 1;

```

Qy      5 YONYVNGEWK--SSVNOIEILSPIDDSLGFVPAMTREEDVDHAKMGREALPAMALTVY 62
Db      7 YKFLYNGEMWRSSSGETIEIPSPYLHEVYIGVQVATRGSEVDALASAKAEQSMVAESLQ 66
Qy      63 ERAQYLHKAADIIIEPKKEIATVLAKEISKAYNASVTEVTRADLIRYAEEGIRLSTSA 122
Db      67 DRAKLYKMADELVMQDEIADIIMKEVGKGYKAKKEVTRADFIIRYTIIEALHMGES 126
Qy      123 DEGGKMDASTGKLAIVIRROPVGIYLAIPYNPVNLSSGKIAPALIGNVMPKRPFG 182
Db      127 MMGDSFPGGTGSKLAIIGRAPLGVLAIPFNNPVNLSSAKLAPALIMGNNAVIFKPAATG 186
Qy      183 SVSGVLAKAFABAGLPAGVFNITTCRGSEIGDYIVHEEENVPINFSTGTPVQGRIGKLA 242
Db      187 AISGIKMVEALHKAQIPKGLVNVATGRSGVIGDYIVHEEGINMVSFTGTTNGKILAKKA 246
Qy      243 GMRPIMLELGGKADGIVLADADLDNNAKOIYAGAYDSGRCCTAIKRYLVVEEVADELAE 302
Db      247 SMIPVLVLELGGKDPGIVREDADLDQDAAHHIVSGAFSYSGRCCTAIKRYLVHEVNADELVG 306
Qy      303 KISENVAKISVGDPEPDNATVTPVIDNNSADFIESTLVNAROKAKELNEFPKDGRLITPG 362
Db      307 LYKEQVAKISVSGSPQDSTIVPLIDDKSADFVQGLVDAVEKAGATVIGNKRENNLIYPT 366
Qy      363 LEDHTLDMKLAEEFPFPIIPIIRYKDAEBAVAIANKSDPGLQSSVFTTRDPQKAFDIAN 422

```

D6	367	LIDHTEEMKAYMEEPFPIILPTIRVSSDECAIETLANKESEFLQASVFYTKINKPAALAN	426
Oy	423	KLEVGTVHNNKYGGRPNPFPLGLKSGAGVOGIRYSIEAMTNKSYIVLDM	474
D6	427	KLETGSVQINGRTERGPDHPFIPIGVKSGMGAGIRKSJLESSTREKVYTLNL	478
RESULT 14			
ID	Q81UL6	PRELIMINARY;	PRT; 479 AA.
AC	Q81UL6; O6IZU6; O6KMW3;		
DT	01-JUN-2003 (TREMBLrel_24,	Created)	
DT	01-JUN-2003 (TREMBLrel_24,	Last sequence update)	
DT	25-OCT-2004 (TREMBLrel_28,	Last annotation update)	
DE	Glyceraldehyde-3-phosphate dehydrogenase, NADP-dependent.		
GN	Name=gpmP; OrderedLocustNames=BA0849, BAS0808, GBA0849;		
OC	Bacillus anthracis.		
OK	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
RN	NCB1_Taxid=1392;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Ames / isolate Porton;		
RX	MEDLINE=22608414; PubMed=127212629; DOI=10.1038/nature01586;		
RA	Read T.D., Peterson S.N., Tourasse N.J., Bailie L.W., Paulsen I.T.,		
RA	Nelson K.E., Tetelien H., Fouts D.E., Eisen J.A., Gill S.R.,		
RA	Holtzapple E.K., Okstad O.A., Helgason E., Risterson U., Wu M.,		
RA	Koloney J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,		
RA	Deboy R.T., Madpu R., Daugherty S.C., Durkin A.S., Hart D.H.,		
RA	Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,		
RA	Benton J.L., Mahmoud Y., Jiang L., Hance I.R., Weidman J.F.,		
RA	Berry K.J., Plant R.D., Wolf A.M., Watkins K.L., Mierman W.C.,		
RA	Hazen A., Cline R.T., Redmond C., Thwaites J.E., White O.,		
RA	Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,		
RA	Hanna P.C., Kolstoe A.-B., Fraser C.M.,		
RT	"The genome sequence of <i>Bacillus anthracis</i> Ames and comparison to		
RL	closely related bacteria."		
RL	Nature 423:81-86(2003).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Ames / isolate 0581;		
RA	Ravel J., Raeko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,		
RA	Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,		
RA	Fraser C.M.;		
RT	" <i>Bacillus anthracis</i> comparative genomics."		
RL	Submitted (MAY-2004) to the EMBL/genbank/DBSJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Sterne;		
RA	Bretlin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,		
RA	Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,		
RA	Richardson P., Rubin E., Rice H.;		
RL	Submitted (JAN-2004) to the EMBL/genbank/DBSJ databases.		
DR	EMBL; AE017026; AAP24851.1;		
DR	EMBL; AE017334; AAT29962.1;		
DR	EMBL; AE017225; AAT5135.1;		
DR	HSSP; Q59931; IECH.		
DR	TIGR; BA0849;		
DR	TIGR; GBA0849;		
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.		
DR	GO; GO:0008152; P:metabolism; IEA.		
DR	InterPro; IPRO02086; Aldehyd_dehydrg.		
DR	Pfam; PF00171; Aldehd; 1.		
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.		
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.		
KM	Complete proteome.		
SC	SEQUENCE 479 AA; 51810 MW; 587029DDB2C8CF652 CRC64;		
Query Match	55.9%; Score 1338; DB 2; Length 479;		
Best Local Similarity	55.9%; Pred. No. 5-e-72;		
Matches	264; Conservative 78; Mismatches 128; Indels 2; Gaps 1		
Oy	5 YONYNGEMWK--SSVNOIEITISPIDDSLSGFVPANTREEVDHAMKAGREALPAAVALTY	62	

Db	7	TKFYINGEMWRESSCETTEIDESPYLEHVIQVQVAILTRGEVDEALIASAKAQSWBASIQ	66
Qy	63	ERAOYLHRAADIIEEDKKEIATVLAKEISKAYNASVTEVTRADILRYAAEGRILSTSA	122
Db	67	DRAKLYKMADELVMQDEIADIIMKEVKGKCAKKEVVRTRADIFIRYIEEALHNGHS	126
Qy	123	DEGKMADSTGHKAVIRRPQVIGVLAIPNYPVNLSSKAPALIGNVNMFKPTGQ	182
Db	127	MMGSPFGGTSKKAIIIRAPLGVVLAIAFPNYPVNLSPAUKLAPALIMNNAVIFKPAIQ	186
Qy	183	SVSGVLAKAPAEAGLPAQVFNTTIGRSEIGDYVEHEHGVNFNTGSTPVGQRIGKLA	242
Db	187	AISGKMWALHKKALPQGLVNVATGRGVSIGDYVEHEGIMNVSFTGGTNGKHLAKQA	246
Qy	243	GMRPIMLELGGKDAIGVLAIDLDNAKQIYAGAYDSGORCTAIRVLVEEVADELAE	302
Db	247	SMILPVLVELGKDRPEIVAEADLDQANHHIYSGASYSQRCCTAIRVLVHENVADELVS	306
Qy	303	KISENVAKLSVGDPPFNATVTPVVIDNSADFTIESLVDAQQAELNEFKEDGRLLTQG	362
Db	307	LVKEQVAKLSVGSPEQDSTVPLVIDDKSADVLQGLVDAVEKGAATVIGNKKERNLIYPT	366
Qy	363	LEPDHYTLDMKLAMEEPPRPILPIIRVKAEEVAALANKSDRELQSSVFRDQKADIAN	422
Db	367	LIDHYTEEMKAVMEEPFPGPILPIIRVSSDEQALIEIANNSEFGLQASVFKDINKKALIAN	426
Qy	423	KLEVGVTHNNKGTGKGNPFPFLGKSGAGVQGIIRYSIEAMTNVSYILDM	474
Db	427	KLETSVQINGRTKRGPDHPFPIGVKSGMGAKGIRKLSIESMTRKVTYVNL	478
RESULT 15			
ID	Q722Z6	PRELIMINARY:	PRT: 477 AA.
AC	Q722Z6:		
DT	05-JUL-2004 (TREMBLrel. 27, Created)		
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)		
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)		
DE	Glyceroldehyde-3-phosphate dehydrogenase, NADP-dependent.		
GN	Ordered locus names=BCB4521.		
OS	Bacillus cereus (strain ATCC 10987).		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=22523;		
RN	[1]		
RP	SEQUENCE FROM N. A.		
RX	PubMed=14960714; DOI=10.1093/nar/gkh258;		
RA	Raebko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,		
RA	Shores K.A., Fouts D.E., Tourasse N.J., Angluoli S.V., Kolonay J.F.,		
RA	Neison W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;		
RT	"The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic		
RT	adaptations and a large plasmid related to Bacillus anthracis pXO1."		
RL	Nucleic Acids Res. 32:977-988(2004).		
DR	EMBL; AE017278; AA543422.1; -.		
DR	HSSP; P56533; 1A4S.		
DR	TIGR; BCE4521; -.		
DR	GO: GO:0016491; F:oxidoreductase activity; IEA.		
DR	GO: GO:0008152; P:metabolism; IEA.		
DR	InterPro: IPR002086; Aldehyd. dehyd.rog.		
DR	Pfam: PF00171; Aldech. 1.		
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.		
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.		
KW	Complete proteome.		
SEQ	SEQUENCE 477 AA; 51638 MW; 0931ID587AE84536 CRC64;		
Query Match 55.8%; Score 1335; DB 2; Length 477;			
Best Local Similarity 55.8%; Pred. NO. 8.3e-72;			
Matches 264; Conservative 78; Mismatches 129; Indels 2; Gaps 1			
Qy	5	YQNVVNGWKK--SSYNQTEILISPIDSSLGFPAMTREVYHAMKAGREALPAMALATY	62
Db	5	YKFYINGEMWRESSCETTEIDESPYLEHVIQVQVAILTRGEVDEALIASAKAQSWBASIQ	64
Qy	63	ERAOYLHRAADIIEEDKKEIATVLAKEISKAYNASVTEVTRADILRYAAEGRILSTSA	122

Db	65	DRÄKTYLKAMDELVNQO:DEIADIIMKEVGKYGKADKKEVVRPADPRTTIEALHHMGES	1245
Qy	123	DEGGRKDA\$IGHLTA\$IRKQPGVITLTA\$PNTY\$PNTLSGSKAPL\$IGNV\$MFRPPTOG	1825
Db	125	MMED\$SPGQ\$K\$K\$T\$A\$T\$IQ\$R\$P\$IG\$V\$T\$A\$P\$NY\$P\$N\$T\$A\$K\$A\$P\$L\$M\$G\$N\$V\$I\$R\$K\$P\$A\$T\$O\$G	1845
Qy	183	\$V\$E\$G\$L\$V\$T\$A\$K\$A\$P\$A\$E\$G\$P\$A\$G\$V\$E\$N\$T\$IT\$G\$E\$S\$E\$D\$Y\$T\$E\$H\$E\$V\$N\$F\$IT\$G\$T\$S\$T\$P\$V\$O\$R\$I\$K\$T\$A	2442
Db	185	A\$I\$G\$K\$I\$K\$V\$E\$H\$A\$H\$Q\$G\$P\$K\$G\$L\$V\$N\$A\$T\$O\$R\$G\$S\$V\$I\$E\$D\$Y\$T\$E\$H\$E\$G\$I\$D\$M\$S\$F\$T\$O\$G\$T\$T\$G\$H\$A\$K\$K\$A	2444
Qy	243	G\$M\$P\$M\$L\$E\$G\$Q\$K\$A\$G\$I\$V\$L\$A\$D\$A\$D\$I\$D\$N\$A\$K\$O\$I\$V\$A\$G\$A\$Y\$D\$S\$G\$O\$R\$T\$A\$K\$R\$V\$L\$Y\$E\$B\$A\$D\$E\$A\$E	3025
Db	245	A\$M\$T\$T\$V\$L\$E\$G\$Q\$K\$P\$G\$I\$V\$R\$E\$D\$A\$D\$Q\$P\$A\$N\$H\$I\$V\$G\$A\$F\$S\$G\$O\$R\$T\$A\$K\$R\$V\$L\$Y\$E\$H\$V\$A\$D\$E\$L\$V\$S	3045
Qy	303	K\$I\$E\$N\$V\$A\$K\$S\$V\$G\$P\$P\$D\$N\$A\$T\$Y\$T\$P\$I\$D\$N\$S\$A\$D\$T\$E\$S\$L\$V\$D\$A\$R\$Q\$K\$A\$E\$L\$N\$E\$P\$F\$O\$R\$G\$L\$T\$P\$G	3625
Db	305	L\$O\$B\$O\$V\$A\$K\$S\$V\$G\$S\$P\$E\$D\$S\$T\$I\$V\$P\$L\$I\$D\$K\$S\$A\$D\$P\$O\$Q\$L\$V\$A\$D\$A\$V\$E\$G\$A\$T\$I\$V\$G\$N\$R\$E\$N\$N\$L\$Y\$P\$T	3645
Qy	422	L\$P\$H\$V\$T\$L\$D\$M\$K\$L\$A\$E\$E\$P\$F\$G\$P\$L\$P\$T\$I\$R\$K\$O\$D\$E\$A\$V\$A\$L\$A\$N\$K\$S\$D\$F\$G\$L\$O\$S\$V\$T\$P\$D\$F\$O\$K\$A\$P\$I\$D\$N	4225
Db	424	L\$I\$D\$H\$V\$E\$D\$M\$V\$A\$E\$E\$P\$F\$G\$P\$L\$P\$T\$I\$R\$S\$D\$E\$O\$A\$E\$L\$A\$N\$S\$E\$B\$E\$L\$O\$A\$S\$V\$T\$K\$O\$I\$N\$K\$A\$P\$A\$T\$A\$N	4224
Qy	475	K\$E\$V\$G\$Y\$V\$H\$I\$N\$K\$T\$G\$R\$P\$D\$N\$F\$P\$E\$G\$L\$K\$G\$S\$A\$G\$V\$Q\$G\$I\$R\$S\$T\$E\$A\$M\$T\$N\$K\$S\$I\$V\$L\$D\$M\$K	4755
Db	477	K\$I\$T\$G\$S\$V\$O\$I\$N\$G\$T\$E\$R\$G\$P\$H\$P\$F\$G\$V\$G\$S\$G\$M\$G\$Q\$G\$I\$R\$K\$S\$T\$E\$M\$T\$R\$E\$K\$V\$V\$N\$F\$K	4775

RESULT 16

ID	081HE6	PRELIMINARY;	PRT;	479 AA.
AC	081HE6;			
DT	01-JUN-2003 (TREMBLrel. 24, Created)			
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase [EC 1.2.1.9].			
DE	1.2.1.9).			
GN	OrderedLocusNames=BC0868;			
OS	Bacillus cereus (strain ATCC 14579 / DSM 31).			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_taxid=226900;			

RP
[1]
RX SEQUENCE FROM N.A.
RX
MELINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
RA
ITANORA N., Shatokin A., Anderson I., Galleron N., Candelson B.,
Kapacetal V., Bhattacharya A., Renik G., Mikhailova N., Lapidus
Chu I., Mazur M., Goldman E., Larsen N., D'Souza M., Walunas T.,
RA
Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
RA
Overbeek R., Kyrpides N.C.;
RT
"Genome sequence of *Bacillus cereus* and comparative analysis with
RT
Bacillus anthracis,"
RL
Nature 423:87-91(2003).
RL
EMBL; AEO17000; AAP07855.1; -.
DR
HSSP; Q59931; IEUH.
DR
GO; GO:0016491; F:oxidoreductase activity; IEA.
DR
GO; GO:0008152; P:metabolism; IEA.
DR
InterPro; IPRO02086; Aldehyde_dehydrog.
DR
Pfam; PF00171; Aldehyd. 1.
DR
PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
DR
PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.
DR
Complete proteome; Oxidoreductase.
QW
SEQUENCE 479 AA; 51801 MW; F49446B919FF8AC34 CRC64;

Query Match 55.7%; Score 1333; DB 2; Length 479;
 Best Local Similarity 55.7%; Pred. No. 11e-71;
 Matches 263; Conservative 79; Mismatches 128; Indels 2; Gaps 1

```

0y 5 YONVUNGEMK -SSVNOIEILSPIDDSGLFVPAMTRREEDVHAKGREALPAWALTVY 62
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 YKFLNCEWMESSGTEILPSPLYLHEVGQVQAITRGVBDALISAKAEQKSNWESLQ 66
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0y 63 ERAQYLKKAADIIRDKKEIATVLAKEISKAYNASVTEVTRADLIRYAEGIRLSTSA 122
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db	67	DRANLYYKWADEIVANNQOEIADIIMKEVGGKQKAKEVARTADPIRTYIEEALHMHGES	1.6
Qy	123	DEGCKRDASTGHKLAVIRBPQVGYVILAAPIIYPPNLSGSLAPALIGGVNVMFKPPTOG	1.82
Db	127	MMDSPPGCTKSKLLAIORAPGLVILAAPPIYPPNLSAAKLAPALIMGNAVIFKPATOG	1.86
Qy	183	SVSGVLAAKAPAEAGLPGVFWPTITNGRSEJIGDYIVHEEBVNFINTFGSPVQORIGKLA	2.42
Db	187	AISGIMVBAHLKAGLPGKLVNVAATGSGSVIGDYLVEHEGIIIMVSTFGTNGTNGHLAKKA	2.46
Qy	243	GMPPIMLTEGGKDAQAGIVLADADLDNAAQIYVAGAVDYSGQRCTAIKTVLVEEVADELAE	3.02
Db	247	SMPLVLEIEGGKDPGLVEKEDADLDQNAHHIYSGAFSISGQRCTAIKTVLVEHENVADLVS	3.06
Qy	303	KISENVAKLSVGPPEFNDAVTVPVIDNSADFTESLIVDARQKAKEINEFKRDLRLLTPG	3.62
Db	307	LTKAQVAELSVGSPBEDSTIVPLIDDKSADVPQGLVDAVEKGAITVIYGNKRENNLIYPT	3.66
Qy	363	LFDHVTLLDKLAMEEPPGFIILPIIRVKDAEEBAVAIANKSDFGLOSSVFTDFQAKAPIIAN	4.22
Db	367	LIDHVTKEKVMVEEPPGFIILPIIRVSDQEQAIETANKSEFGLOASVFTDKINKAFAIAN	4.26
Qy	423	KLEVGIVAHINKTGRGPDNFPPLGLKSGAGVQGRISIEAMTVKSIIVLDM	4.74
Db	427	KIEFGSVQINGRIERGPDHFPPIIYKSGGMAQGRISLEMTREKTVANL	4.78

RESULT 17

08XHP4		
ID	08XHP4	PRELIMINARY; PRT; 482 AA.
AC	08XHP4	
DT	01-MAR-2002 (TrEMBLrel. 20, Created)	
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)	
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE	MADP-dependent glyceraldehyde-3-phosphate dehydrogenas.	
GN	Name=gadp; Ordered locus Names=CPE2138;	
OS	Clostridium perfringens.	
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;	
CC	Clostridium.	
NCBI_TaxID=1502;		

RN SEQUENCE FROM N.A.
 RP [1]
 RC MEDLINE=21664373; PubMed=11792842; DOI=10.1073/pnas.0224937999
 RX Shlimizu T., Ohtani K., Hirakawa H., Oshshima K., Yamaashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kihara S., Haysashi H.;
 RT "Complete genome sequence of *Clostridium perfringens* strain
 RT flesh-eater";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:396-1001(2002).
 DR EMBL; AP003194; BAB82144.1; -.
 DR HSSP; Q59931; 1EUG.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR Pfam; PF001771; Aldehyd_1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYC; UNKNOWN_1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 SC Complete proteome.
 SQ SEQUENCE 482 AA; 52763 MW; 2DD7DA4A3BC03EC4 CRC64;

Query Match	55.3%	Score 1323;	DB 2;	Length 482;
Best Local Similarity	52.3%;	Pred. No. 4.4e-71;		
Matches 251; Conservative	99;	Mismatches 124;	Indels 6;	Gaps 3

Oy 2 TKE--YONYNGEM--KSSVNOIEILSEIDDSLSGFPAMTREVEDHAMKA GREALPAMA 57
 ||| : :: | : | | | | : : : : : | : : :
Db 3 TKEIIFNFMWGANVEENKSDKLIDI KSLDSSLGRLOSLEKEDVDKTIIDNAEKQXSWN 62

Oy	58	A L T V Y B R A Q Y L H K A A D I I E R D K E E I A T V L A K E I S K A N V A S T E V V T A D L I R V A A E G I R	117
Dd	63	E V P L N E A H Y L Y T T A D L E E H K E E L A M I R E V A D O K S F A L S E V S R P A D I T R F A D A K S	122
Oy	118	L T S A D E G K M D A S T G H K L A V I R P O V G I V A I A P Y N P V N L I S G S K I A P A L I G N V V M F K	177

```

Db      123 MDGESIPDNFPGFNRSKISVTRREPLGVTLAISPFNYVNLASAKIAPALVAGNSVLMK 182
Qy      178 PPTQGSVSGVLAKAPAEAGLPAGVNTTITGRSGSEIGDYIVHEEVENFINFTGISTVQGR 237
Db      183 PATQGSLSALFLAKMFDAGLPQGVNTITVGRSGSEIGDYCVTDKIDFINFTGSTEVGR 242
Qy      238 IGLTAGRPIMLELGGKDGIVLADADLDMAAQIYAVGAVDYGSGORTAIKRYLVYEEVA 297
Db      243 ISRTIWMKPLMELGGDAIVLEDDLDLTNNITIAIGYSISGORTAIKRYLVYEEVA 302
Qy      298 DELAEKISENVAKLSVGDPF--DNATVTPVIDNSADFIESTLVVDAROKGAKELNEFKD 355
Db      303 DKIVKELKEKISLTKNGNPLFEVEDVTIIVPIDMKADFWVELIDDAKEKATLVCGGHE 362
Qy      356 GRLLTPGLFDHVTLMKLANEPPPGILPIIRYKDAEBAVALANKSDPGLQSSVTRDPQ 415
Db      363 GNIEATLFDNVTITDRMLAEEPPGVLPIIRINDDEAIKIANASKYGLQSSVTRDIN 422
Qy      416 KAPDIANKLEVGVHNNKTRGPDNPPLGLKSGAGVGGIRYSIEAMTNVKSIVLDMK 475
Db      423 EAFYIANKLEVGVHNNKTERGPDHPPLGVKSGMGVQGIKYSIEAMSRPATVITIR 482

```

RESULT 18

```

O9KAO0      PRELIMINARY;      PRT;      481 AA.

```

```

AC      O9KAO0;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE      NADP-dependent glyceralddehyde-3-phosphate dehydrogenase (EC
1.2.1.9).
GN      Name=gapN; OrderedLocusNames=BH2237;
OC      Bacillus halodurans.
OC      Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX      NCBI_TaxID=86665;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C-125;
RX      MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA      Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA      Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA      Horikoshi K.;
RT      "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT      halodurans and genomic sequence comparison with Bacillus subtilis."
RL      Nucleic Acids Res. 28:4317-4331(2000).
RL      EMBL: AP001514; BAB05956.1; -.
RL      PIR: E83929; E83929.
DR      HSSP: Q59931; IEUH.
DR      GO: GO:0016491; F:oxidoreductase activity; IEA.
DR      GO: GO:0008152; P:metabolism; IEA.
DR      InterPro: IPR002086; Aldehyd_dehydrog.
DR      Pfam: PF00171; Aldehyd. 1.
DR      PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
DR      PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW      Complete proteome; Oxidoreductase.
SQ      SEQUENCE 481 AA; 51452 MW; 82769AC55C95F0B9 CRC64;

```

Query Match 52.0%; Score 1244; DB 2; Length 481;

Best Local Similarity 51.4%; Pred. No. 2.3e-66;

Matches 240; Conservative 95; Mismatches 130; Indels 2; Gaps 1;

```

Qy      10 NGEWKS--VNOIILSPIDDSILGFVPAMTRREVDHMAKREALPAMALTVYERAOY 67
Db      15 NGMWESRTGERISISAPAGVALGISIPALSGEVDALIGADAQKIRIRIHERVDL 74
Qy      68 LHAADAIIEEDKEIATVLAKEISKAYNAVTEVVTADILRYAAEGRIILSSADEGK 127
Db      75 LYVADLIEREKIIIGELIWEVAKPKKSIIGVSRADIRITADBAKLNETTLKGDQ 134
Qy      128 MDASTGKTLAVIRROPVGLVATAPYNYVNLGSKIAPALIGNVVMEKPPQGSVSG 187
Db      135 FKGGSSKTLAVREPLGVVLAISPFNYVNLAAKIAPALVYGNVVFPEPAIQSLSGI 194

```

```

Qy      188 VLAKAPAEAGLPAGVNTTITGRSGSEIGDYIVHEEVENFINFTGISTVQGR 247
Db      195 KVEEALADGABEGITIQVTTGRGSYIGDLVHPGIDMTTFGGITGRISKAMMIV 254
Qy      248 MLELGGKDGIVLADADLDMAAQIYAVGAVDYGSGORTAIKRYLVYEEVADEIAEKISEN 307
Db      255 VLELGGKDPALVDDADLKLTSQIVSAGFSISGQRCCTAIKRYVODSVADQULVNIKEL 314
Qy      308 VAKLSVGDPPDNATVTPVIDNSADFIESTLVVDAROKGAKELNEFRDRLITPLGFHY 367
Db      315 VEQLTFGSDEDDADITPVIDEKSAFTIQLIDDALENGATLLSGNRQGNLSPILLDDV 374
Qy      368 TIDMKLANEPPPGILPIIRYKDAEBAVALANKSDPGLQSSVTRDPQAFIDIANLEFG 427
Db      375 TPAMRYAMEPPPGVLPIIRYKDAEBAVALANKSDPGLQSSVTRDPQAFIDIANLEFG 434
Qy      428 TVHINNKTGRGPDNPPLGLKSGAGVGGIRYSIEAMTNVKSIVLDM 474
Db      435 TVHINAKTERGPDHPPLGVKSGGLGVQGIKFSLSMTIRRYTVLNL 481

```

RESULT 19

```

O9SNX8      PRELIMINARY;      PRT;      496 AA.

```

```

AC      O9SNX8;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE      Nonreversible glyceralddehyde-3-phosphate dehydrogenase (EC
1.2.1.9).
OS      Apium graveolens (Celery).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC      campanulids; Apiales; Apiceae; Apioideae; apioid superclade;
OC      Apium clade; Apium.
OX      NCBI_TaxID=4045;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=leaf;
RX      MEDLINE=20438225; PubMed=10982446; DOI=10.1104/pp.124.1.321;
RA      Gao Z., Loeschner W.H.;
RT      "NADPH supply and mannitol biosynthesis. Characterization, cloning,
RT      and regulation of the non-reversible glyceralddehyde-3-phosphate
RT      dehydrogenase in celery leaves."
RL      Plant Physiol. 124:321-330(2000).
RL      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=leaf;
RA      Zhitang G., Loeschner W.H.;
RA      Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF196292; AAF08296.1; -.
DR      HSSP: O59931; IEUH.
DR      GO: GO:0016491; F:oxidoreductase activity; IEA.
DR      GO: GO:0008152; P:metabolism; IEA.
DR      InterPro: IPR002086; Aldehyd_dehydrog.
DR      Pfam: PF00171; Aldehyd. 1.
DR      PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
DR      PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW      Oxidoreductase.
SQ      SEQUENCE 496 AA; 53173 MW; D86B4C01E0A3642 CRC64;

```

Query Match 49.0%; Score 1171.5; DB 2; Length 496;

Best Local Similarity 50.3%; Pred. No. 5.3e-62;

Matches 240; Conservative 83; Mismatches 143; Indels 11; Gaps 5;

```

Qy      5 YQNVNNGEKK--SYVNOIILSPIDDSILGFVPAMTRREVDHMAKREALPAMALTVY 62
Db      16 FYTISYGEKWKSSGKSVAILINPTTRMTQFOACQESVNNKAMEYAKKVOQWAKTPLM 75
Qy      63 EBAQYLHKAADIIEEDKEIATVLAKEISKAYNAVTEVVTADILRYAAEGRIILSSADEG 122
Db      76 KBAELHLKAAALILKEKKAIALDCLVKEIAKPAKDSVTEVVRSGDLVSYCAEGBVRL--L 132

```


QY 123 DEGGM--DASTGH---KLAIVIRQPVGIVLAIAIPYNPVNLGSKIAPALIGNVVMEK 177
 DB 123 GEGKFLVDSPPGNERTKYCLTSKIPGLVITAIAPPNTYVNLAVSKIGPALLIAGNALVLK 192
 QY 178 PPTQSSVGLVLAFAKFAEAGLPAQVNTITGSGSEIGDYIVHEEYVNFNFPGSPVQGR 237
 DB 193 PPTQAVVACLHMHVCFHLAEPFKGLISCTTGKSGEIGDITMHPGNCISFPGG--DTGIA 251
 QY 238 IGLKAGMRPIMLELGGKADGIVLADADLDNAKQIVAGAYDYSQGRCTAIRVLYVEVA 297
 DB 252 ISKKGAMVPLQWELGCKDACIVLEBDADLVASNTYKGSFSGQACHTAIKILVWQSYA 311
 QY 298 DELAKISENVAKLSVGPDPNATVTPVIDNSADFIESLVDYDARQKAKELNEFRDGR 357
 DB 312 DTLVEKVAKAKLTVGPEDNSDITPVVSESSANFIEGLVDAKEKATPCQETKREGN 371
 QY 358 LITPGLPHTVILDMKLAMEBPPGPIIPITIRVDAEVAVALANKSDPGLQSSVFTPDFOKA 417
 DB 372 LITPGLLDVVKPDMRIAMEBPPGPIIPVIRINSABEGIHNCNANFGLQGVFTPDINKA 431
 QY 418 FDIANKLEVGTIVHINKTKGRGPNPPLGLKSGAGVQIGIRYSIEAMTNVKSIVLDM 474
 DB 432 MLISDAMESGTTQINSAPARKGDHPFPQGLKDSGIGSGQITMNSINMTIKITIVNL 488

RESULT 20
 Q9ZUG8 PRELIMINARY; PRT; 496 AA.

AC Q9ZUG8, 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative NADP-dependent glyceroldehyde-3-phosphate dehydrogenase
 (ACG24270/F2J2D4.18).
 GN Name=ACG24270;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
 RA Shen M., Rensing C.M., Fraser C.M., Somerville C.R., Venter J.C.,
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Town C.D., Kaul S.,
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chan H., Kim C.J., Koesema E., Meyers M.C., Shim P.,
 RA Treacy S.E., Banu U., Bowser L., Carninci P., Chung M.K.,
 RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam J., Lee J.M., Lin J., Liu S.X.,
 RA Marindun M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,
 RA Sakano H., Saiturai T., Satou M., Seki M., Southwick A., Tang C.C.,
 RA Toriumi M., Yamada K., Yu G., Shinzaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.,
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
 RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
 RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,
 RA Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Shim P., Tang C.C., Toriumi M., Walleander E.K., Wong C.,
 RA Wu H.C., Yamada K., Yu G., Yuan S., Shinzaki K., Ecker J.,
 RA Theologis A., Davis R.W.,
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Shim P., Bowser L., Carninci P.,
 RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Marindun M., Narusaka M., Nguyen M., Onodera C.S.,
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
 RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.,
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 RA EMBL; AC005967; AAD03188.1; -
 DR EMBL; AY037205; AAK59790.1; -
 DR EMBL; AY136409; AAM97075.1; -
 DR EMBL; BT004551; AAO42797.1; -
 DR PIR; F84634; F84634.
 DR HSSP; Q59931; 1E0H.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR02086; Aldehyde_dehydratase.
 DR Pfam; PF00171; Aldehdh; 1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.
 SQ SEQUENCE 496 AA; 53060 MW; 9231656A951175D5 CRC64;

Query Match 48.8%; Score 1168.5; DB 2; Length 496;
 Best Local Similarity 50.3%; Freq. NO. 8e-62; Mismatches 142; Indels 11; Gaps 5;
 Matches 240; Conservative 84; Mismatches 142; Indels 11; Gaps 5;

QY 5 YQNVNGEMK--SSVNOIEILSPIDSSLGVPVMTREEDVHAMKAGREALPAMALTVY 62
 DB 16 YKYVADGEMTKSSSSKSAIIMNPATKTYQVQACTOEVAIVAMELAKSQAQSKAKTFLW 75
 QY 63 ERAQYLHRAADIIEKKEIAITVLAKEISKAYNMSVTEVATDILIRYABEGIRLSTSA 122
 DB 76 KRAELIHRAAAILKONKAPMAESLYKEIAKPAKOSVTEVNSGDLISYCAEBGVRI--L 132
 QY 123 DEGGM--DASTGH---KLAIVIRQPVGIVLAIAIPYNPVNLGSKIAPALIGNVVMEK 177
 DB 133 GEGKFLVDSPPGNDRTRYCLTSKIPGLVITAIAPPNTYVNLAVSKIGPALLIAGNALVLK 192
 QY 178 PPTQSSVGLVLAFAKFAEAGLPAQVNTITGSGSEIGDYIVHEEYVNFNFPGSPVQGR 237
 DB 193 PPTQAVVACLHMHVCFHLAEPFKGLISCTTGKSGEIGDITMHPGNCISFPGG--DTGIA 251
 QY 238 IGLKAGMRPIMLELGGKADGIVLADADLDNAKQIVAGAYDYSQGRCTAIRVLYVEVA 297
 DB 252 ISKKGAMVPLQWELGCKDACIVLEBDADLVASNTYKGSFSGQACHTAIKILVWQSYA 311
 QY 298 DELAKISENVAKLSVGPDPNATVTPVIDNSADFIESLVDYDARQKAKELNEFRDGR 357
 DB 312 DTLVEKVAKAKLTVGPEDNSDITPVVSESSANFIEGLVDAKEKATPCQETKREGN 371
 QY 358 LITPGLPHTVILDMKLAMEBPPGPIIPITIRVDAEVAVALANKSDPGLQSSVFTPDFOKA 417
 DB 372 LITPGLLDVVKPDMRIAMEBPPGPIIPVIRINSABEGIHNCNANFGLQGVFTPDINKA 431
 QY 418 FDIANKLEVGTIVHINKTKGRGPNPPLGLKSGAGVQIGIRYSIEAMTNVKSIVLDM 474
 DB 432 MLISDAMESGTTQINSAPARKGDHPFPQGLKDSGIGSGQITMNSINMTIKITIVNL 488

RESULT 21
 GAPN MAIZE STANDARD; PRT; 498 AA.
 ID G4372;
 AC G4372;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE NADP-dependent glyceroldehyde-3-phosphate dehydrogenase (EC 1.2.1.9)
 DE (Non-phosphorylating glyceroldehyde 3-phosphate dehydrogenase)
 DE (Glyceroldehyde-3-phosphate dehydrogenase [NADP+]) (Triosephosphate
 dehydrogenase).
 GN Name=GPNI;
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC	Spermatophyta: Magnoliophyta, Liliopsida: Poales, Poaceae;
OC	PACCAD clade, Panicoideae; Andropogoneae; Zea.
OX	NCBI_TaxID=4577;
RP	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=cv. KM5330, TISSUE=shoot;
RX	MEDLINE=941803877, PubMed=7545914;
RA	Hadenicht A., Hellman U., Cerif R.;
RT	"Non-phosphorylating GAPDH of higher plants is a member of the
RT	aldehyde dehydrogenase superfamily with no sequence homology to
RT	phosphorylating GAPDH.";
RU	J. Mol. Biol. 237:165-171(1994).
CC	-1- FUNCTION: Important as a means of generating NADPH for
CC	biochemical reactions.
CC	-1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + NADP(+) + H(2)O
CC	= 3-phospho-D-glycerate + NADPH.
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.
CC	-1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; X75326; CAA53075.1; .
DR	PIR; S43833; S43833.
DR	HSSP; Q59931; 1E0H.
DR	MaizeDB; 78926; .
DR	InterPro; IPR002086; Aldehyde_dehydr.
DR	Pfam; PF00171; Aldehd; 1.
DR	TIGRFAMs; TIGR01804; BADH; 1.
DR	TIGRFAMs; TIGR01722; MMSDH; 1.
DR	TIGRFAMs; TIGR01780; SSADH; 1.
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW	NADP; Oxidoreductase.
FT	NP BIND 247 251 NAD (ADP part) (By similarity).
FT	ACT_SITE 266 266 By similarity.
FT	ACT_SITE 300 300 By similarity.
FT	SEQUENCE 498 AA; 53146 MW; 7AF1CD0CABAEE39 CRC64;
Query Match	48.7%; Score 1164.5; DB 1; Length 498;
Best Local Similarity	49.0%; Pred. No. 1.4e-61;
Matches 234; Conservative	88; Mismatches 143; Indels 13; Gaps 5;
OY	5 YONYNGEWSKSVN--QIEILSPIDDSLSGLFVPMTREBEVDHANKAGREALPAAALTVY 62
DB	18 YRYVADGEMRISASGSKVAIYNPTTRKIQYRVQCTGSEVNRKADDAKVAQKAMARPLW 77
OY	63 ERAQYLKKAADIIEBDKEELATVLAKESIKAYNNSVTEVPTADLIRYAEERGLSTSA 122
DB	78 KRADVLHKAAILKHKHAKPIAECLVKEIAKPAKVAESEVNSGDLVSYTAEGRILGE- 136
OY	123 DEGGMM---DASTG---HKLVIRQPVGIATLAPYVNVNLSGSKAPALIGQVYMF 176
DB	137 ---GLVVSDFPGRERNKYCLSSKTIPLGVVLAIPFVNPANLGSKIGPLLIGNMLVL 193
OY	177 KPPTGVSGLVLAKAFAEAGLPAGVFNTITGRGSEIDYIVHEBEVNFIFGTSTPVQ 236
DB	194 KPPTGVAVALHMVHCFILAFPGKGLISCVGKSEIDPLVHMPGVNCLISFTGG-DTGI 252
OY	237 RIQKLAKRPIMLLEIGKQDAGIVLADADLDNAKOIVAGAYDYGQRTAIKRYLVVEV 296
DB	253 AISKAGGVPLQMEIGSGDACLVEDADLDIVSNATVIGSGSYSGQRCTAKVVLIMESI 312
OY	297 ADELAETKISENVAKLSVGDPPDNVATVVIDNNSDPIESLIVVAROKGAKEINFFRQD 356
DB	313 ADAYVQAKNAKIAKLVKQVPEDDSITIVVESSANFLIEGLVMDKEGALFCCOETREB 372
OY	357 RLTPGGLDHTVLTDMKLAEEBFGELIDPIIRVKQAEAAVIAANKSDPGLQSGVETTRFOK 416

ID	Accession	Species	Protein Name	Length	Score	Pred. No.	Indels	Gaps
Db	373	NLWPLLLDHRPDRIRIAMEEPFGVLPVIRINSVVEEGHHCNANFGLQCTFTDINK		432				
Oy	417	AFDINKLEVGTVINHNKTRGPDNPFPFGLKSGAGVQGIHYSIAMTNVKSIVDM		474				
Db	433	AILISDAIETGTVQINSAPARPDHFSFQGLKDSIGSGQITNSIMMKVYSTVINL		490				
RESULT 22								
	ID	GAPV_PEA	STANDARD:	PRT:	496 AA.			
AC	P81406							
DT	15-DEC-1998	(Rel. 37, Created)						
DT	15-DEC-1998	(Rel. 37, Last sequence update)						
DE	05-JUL-2004	(Rel. 44, Last annotation update)						
DE		NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9)						
DE		(Non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase)						
DE		(Glyceraldehyde-3-phosphate dehydrogenase [NADP+]) (Triosephosphate dehydrogenase).						
DN		Name=GAPN;						
OS		Pisum sativum (Garden pea).						
OC		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
OC		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;						
OC		eurosid 1; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.						
OX		NCBI_TaxID=3888;						
RP		[1]						
RT		SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.						
RC		STRAIN=cv. Rosakrone; TISSUE=shoot;						
RA		MEDLINE=94180387; Pubmed=7545914;						
RA		Hadenicht A., Hellman U., Cerif R.;						
RT		"Non-phosphorylating GAPDH of higher plants is a member of the						
RT		aldehyde dehydrogenase superfamily with no sequence homology to						
RT		phosphorylating GAPDH.";						
RL		J. Mol. Biol. 237:165-171(1994).						
CC		-I- FUNCTION: Important as a means of generating NADPH for						
CC		biosynthetic reactions.						
CC		-I- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + NADP(+) + H(2)O						
CC		= 3-phospho-D-glycerate + NADPH.						
CC		-I- SUBCELLULAR LOCATION: Cytoplasmic.						
CC		-I- SIMILARITY: Belongs to the aldehyde dehydrogenase family.						
CC		-----						
CC		This SWISS-PROT entry is copyright. It is produced through a collaboratio						
CC		on between the Swiss Institute of Bioinformatics and the EMBL outstation						
CC		at the European Bioinformatics Institute. There are no restrictions on its						
CC		use by non-profit institutions as long as its content is in no way						
CC		modified and this statement is not removed. Usage by and for commercia						
CC		entities requires a license agreement (See http://www.isb-sib.ch/announce						
CC		or send an email to license@isb-sib.ch).						
CC		-----						
DR	EMBL: X75327; CAA53076.1; -.							
DR	HSBP; O59931; IEDH.							
DR	InterPro: IPR002086; Aldehyde_dehydr.							
DR	PIfam; PF00177; Aldedh; 1.							
DR	TIGRFAMs; TIGR01804; BADH; 1.							
DR	TIGRFAMs; TIGR01722; MMSDH; 1.							
DR	TIGRFAMs; TIGR01780; SSADH; 1.							
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.							
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.							
KW	Direct protein sequencing; NADP; Oxidoreductase.							
FT	NP BIND 245 249							
FT	ACT SITE 264 264							
FT	ACT SITE 264 264							
FT	ACT SITE 298 298							
FT	ACT SITE 298 298							
SO	SEQUENCE 496 AA; 53004 MW; 52PADDC2F77C613FA CR664;							
Query Match 48.5%; Score 1161.5; DB 1; Length 496;								
Best Local Similarity 49.5%; Pred. No. 2.1e-61;								
Matches 236; Conservative 85; Mismatches 145; Indels 11; Gaps 5								
Oy	YONYNNGEMKSSVN-OEIIISPDISSLGFPVPMTRVEVDHAMKRGREALPMAALTGY 62							
Db	16 YKYVADGEMKSTSKSVAILNPTTRKQYKQAQSGEIVNKVMSAKAQSMAATPIW 75							
Oy	63 ERAQYIHR							

```

Db      76 KRAELHKAALIKHEKKAIAECLVKEIKAPKADAVTEVRSGLVSCAEGRVRL---L 132
Qy      123 DGGGCK--DASTGH---KLAVIRROVGVIVLAIPNYPVNLGSGKIAALLGANNVMMK 177
Db      133 GGGKFLVSSPFPENERTKYCLTSKIPLVGIVLAIPPNYPVNLAVSKIAALLAGNSIVLK 192
Qy      178 PPTQGSVSGLVAKAPAEAGLPAQVFNITGTGSGSEIGDYIVHEEENFNFNGSTFVQR 237
Db      193 PPTQGAVALAHVHCHLAGFPKGLISCTYGKSEIGDFTMHGPNCSFTGQ--DTGIA 251
Qy      238 ICKLAGMRDIMELEGKDGAGIVLADADLNNAKQIVAGAYDYGSGRCTAIKRVLVEEVA 297
Db      252 IKSXGMIPILOMELGKDGACIVLEADLDLVANIIKGGFSYSGQRCTAVKVLVMEVSA 311
Qy      298 DELAEKISENVAKLSTGDPFDNATVTPVLDNSADTIESLVVDAROKKAELEPFKDR 357
Db      312 DALVEKVKYKAVKLSTGPEPDSITTPVVSSESNFIEGLVMDAKKGAFTCOEYKREGN 371
Qy      358 LITPGLFDHYTLDMKLAKEEPFGPIPIIRVKDAEVAVALANKSPGLQSSVFTRDFOK 417
Db      372 LTIWPLLDVHRPDMRIAMEEPFGVLPVITINSVEEGIHHCNANSFGLQGVFTKDINKA 431
Qy      418 FDIANKLEVGTVHINNKTRGPDNFPFLGLKSGSAGVQIRYSIEAMTVKSYIVLDM 474
Db      432 IMISDAMESGTVOINSAPARPGDPHFPFGIKDSIGSGQITNSINMTTKVSTVINL 488

RESULT 23
Q629G0 PRELIMINARY; PRT; 499 AA.
ID 0629G0
AC 0629G0;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
GN Name=P0528B09.18; Synonym=P0524F03.45;
OS Oryza sativa (Japanese cultivated group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004703; BAD09910.1; -.
DR EMBL; AF004702; BAD09896.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IP002086; Aldehyd_dehydrcog.
DR Pfam; PF00171; Aldehyd. 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN.1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN.1.
SQ SEQUENCE 499 AA; 53369 MW; 2477C6DFB034DB34 CRC64;

Query Match 48.4%; Score 1158.5; DB 2; Length 499;
Best Local Similarity 49.0%; Pred. No. 3.2e-61;
Matches 224; Conservative 87; Mismatches 144; Indels 13; Gaps 5;
Qy 5 YONYVNGEKKSVN--QIRILSPIDSSIGFVPAATREEDHAMKAGREALPAMALTVY 62
Db 19 YRYVADGEWRVSSAGSVAIVNTTTLQRYOACQOEENKMKMETAKYAQKAMATPIWM 78
Qy 63 ERAQVYHKAADIERDKEEIAVLAKESKAYNAVSTEVVTRADLIRYABEGIRLSTSA 122
Db 79 KRAELHKAALIKHEKTPIAECLVKEIKAPKADAISEVRSGLVSYAESEGVRLIG- 137
Qy 123 DGGGCK---DASTG---HKLAVIRROGVIVLAIPNYPVNLGSGKIPALIGVNMVF 176
Db 138 ---GKLLVSDSPFGENRNKYCLSSKVPGLGVLAIPPNYPVNLAVSKIAALLAGNALVL 194
Qy 177 KPTQGSVSGLVAKAPAEAGLPAQVFNITGTGSGSEIGDYIVHEEENFNFNGSTFVQR 236

```

```

Db      195 KPTQGAVALAHVHCHLAGFPKGLINCVKSGSEIGDFTMHGPNVCISFTGQ--DTGI 253
Qy      237 RIGKLAGMRPIMELEGKDGAGIVLADADLNNAKQIVAGAYDYGSGRCTAIKRVLVEEV 296
Db      254 AISKAGVAVPILOMELGKDGACVLEADLDLVANIVAGSGSRCTAVKVLVIMESV 313
Qy      297 ADELAEKISENVAKLSTGDPFDNATVTPVLDNSADTIESLVVDAROKKAELEPFKDR 356
Db      314 ADIVEKVKAKIAKLTVPPEADSDITTPVVSSESNFIEGLVMDAKKGAFTCOEYKREG 373
Qy      357 RLTTPGLFDHYTLDMKLAKEEPFGPIPIIRVKDAEVAVALANKSPGLQSSVFTRDFOK 416
Db      374 NLTIWPLLDVHRPDMRIAMEEPFGVLPVITINSVEEGIHHCNANSFGLQGVFTKDINK 433
Qy      417 AFDIANKLEVGTVHINNKTRGPDNFPFLGLKSGSAGVQIRYSIEAMTVKSYIVLDM 474
Db      434 IMISDAMESGTVOINSAPARPGDPHFPFGIKDSIGSGQITNSINMTTKVSTVINL 491

RESULT 24
GAPN NICPL
ID GAPN NICPL
AC P93338;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9)
DE (Non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase)
DE (Glyceraldehyde-3-phosphate dehydrogenase [NADP+]) (triosephosphate
DE dehydrogenase).
GN Name=GAPN;
OS Nicotiana glumabiginifolia (Leadwort-leaved tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4092;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98036054; PubMed=9370287; DOI=10.1016/S0378-1119(97)00320-X;
RA Habenicht A., Quesada A., Cerit R.;
RT "Sequence of the non-phosphorylating glyceraldehyde-3-phosphate
RT dehydrogenase from Nicotiana glumabiginifolia and phylogenetic origin
RT of the gene family.";
RL Gene 198:237-243(1997).
CC -!- FUNCTION: Important as a means of generating NADPH for
CC biosynthetic reactions.
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + NADP(+) + H(2)O
CC = 3-phospho-D-glycerate + NADPH.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U87848; AAB47571.1; -.
DR HSSP; O59931; IEH.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; Aldehyd. 1.
DR TIGRFAMs; TIGR01804; BADH. 1.
DR TIGRFAMs; TIGR01722; MMSDH. 1.
DR TIGRFAMs; TIGR01780; SSADH. 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
DR NADP; Oxidoreductase.
FT NP BIND 245 249 NAD (ADP part) (By similarity).
FT ACT SITE 264 264 By similarity.
FT ACT_SITE 298 298 By similarity.

```

SEQ	SEQUENCE	496 AA;	53143 MW;	6F350F1688D7A108 CRC64;
Query Match	48.3%;	Score 1155.5;	DB 1;	Length 496;
Best Local Similarity	49.5%;	Pred. No. 4,8e-61;		
Matches	236;	Conservative	86;	Mismatches 144; Indels 11; Gaps 5;
OY	5 YQNVNGEKSSVN--QIEIISPIDSSLGFPVPMTRTEEDHAMKAGEBALPMAALTVY	62		
DB	16 FRYSEGEEMKKSAGSKVAIINPTTRKTYQKVQACTOGEVNVKMEVALTAQGSMAKTPLM	75		
OY	63 ERAQVLHAAADIIEEDKEEIALAVLAKETISKVAMNVTVEVPRADILIRAAEEGRILTSIA	122		
DB	76 KRAELHRAAALLKSHKAPLAECLVKEILAKPKADQVTEVNSGDLVSYTAEGVNI--L	132		
OY	123 DEGGNM--DASTGH--KLAVIRROPGVIGLAIAPYNPVNLGSGKIPALIGGVNWFK	177		
DB	133 GEGRKLVSDSPFGNERTKCYCLTSKPIPLGVITAIIPFPNYPNVLAWSKIPALIAAGSLVK	192		
OY	178 PPTQGSVSGVLAKAPAEAGLPAGVFNITGSGSEIGDYVEHEEVNFINTGSPVGOR	237		
DB	193 PPTQGAVALCLMHVHCFHLIAGPFGKILISCVTGGSGSIEGDFLTMHPVGHISFTGG--DTGVA	251		
OY	238 IGLKAGRPIMLELGGKQAGVITLADADIDNAAKQIVAGAYDYSGORCTAIRVLYVEEVA	297		
DB	252 ISKRKGMPLQWELGGKQACIVLEADADIDLAAGSIVKGSFGFSYSGORCTAIRVLYMESVA	311		
OY	298 DELAEKISEBNVAKLSVGDPEDNATVPIYIDNSADPITSEIIVDAROKAKELNEFRKDRK	357		
DB	312 DALVKKVNAKAKVLTVGPREDCCDITPVVSSNAFIEGLVWDAAQKATTCQOYKREGN	371		
OY	358 LITPGLFPHVTLDMKLAMEEPPGPILPTIRVYDAEAAVALANKSPDGLQSSVFTDRFOKA	417		
DB	372 LIMPILLNVNRPDMRIAMEEPPGPVLPYIRINSVEGIIHHCNASSFGLQGCVPFTDINKA	431		
OY	418 FDIANKLEVGTVHINNTKGRGPDNFPFLTGLKSGAGVQGISIRYSIEAMTVNVSIVLDM	474		
DB	432 ILISAMETGTVQIINSAPARQDHPFGIIXDSIGISOGITNSIMMTKXKTVINTL	488		
RESULT	25			
ID	08S4Y9	PRELIMINARY;	PRT;	507 AA.
AC	08S4Y9;			
DT	01-JUN-2002 (Tremblrel. 21, Created)			
DT	01-JUN-2002 (Tremblrel. 21, Last sequence update)			
DE	01-MAR-2004 (Tremblrel. 26, Last annotation update)			
DE	Glyceralddehyde-3-phosphate dehydrogenase.			
OS	Oryza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzae; Oryza.			
DX	NCBI_TaxID=4530;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	PubMed=12582559;			
RA	Armugam Pillai M., Lihuang Z., Akiyama T.;			
RT	"Molecular cloning, characterization, expression and chromosomal			
RT	location of OsGADH, a submergence responsive gene in rice (Oryza			
RT	sativa L.).";			
RL	Theor. Appl. Genet. 105:34-42(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Pillai A., Akiyama T.;			
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF357884; AAM00227.1; --.			
DR	HSSP; Q59931; IEUH.			
DR	Graeme; 08S4Y9; --.			
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.			
DR	GO; GO:0008152; P:metabolism; IEA.			
DR	InterPro; IPR002086; Aldenhyd_dehydrog.			
DR	Pfam; PF00171; Alded; 1.			
DR	PROSITE; PS00070; ALDEHYDE DEHYDR GLY; UNKNOWN 1.			
DR	PROSITE; PS00687; ALDEHYDE DEHYDR GLY; UNKNOWN 1.			

Query Match	47.9%	Score 1145.5	DB 2	Length 507
Best Local Similarity	48.4%	Pred. No. 2e-60		
Matches 235	Conservative 88	Mismatches 142	Indels 21	Gaps 7
5 YONYVNGKSSV-N-QIETLSPIDSSLGFPVPMTRTEEDVHANKAGREALPAMALTVY	62			
19 YRYVADGEMVRSASGSAVAIVNPTTRLTQYVQACTOEKVNKMETALVQAOKAMARPTLM	78			
63 ERAQVILHQAADIIITRDKKEINTVLAKETSKAYNASVTEVVTADLLIRPAAEGRILSTSA	122			
79 KRAELIHKAAALIKHEKTPIAECVKEITAKPAKDAISEVNSGDLVSTYAEGRVILIGE	137			
123 DEGGKM--DASTG--HKLAVIRROPYGIYLAIAFYNYPVNLSGSKTAPALIGNVYMF	176			
138 ---GKLVSDBFPNGNERKTKLSSKVPGLGVALLPFPYVNNLAVSKIGPALLIAGNALVL	194			
177 KPPTGVSGLVTLAKAFAPAGLPAQVNTITIGRSEIGDIYVEHEVNFIFNGSTPYQO	236			
195 KPPTGVAVALHMHVCFHLPAGFPGGLINCVTGKSEIGDFLTMHPRGVNCISFTGG-DTGI	253			
237 RIGKLACGRPMLLEGGKDAQIVLADADLDNAAQIYAGAYDYGQRTAKRLVYAEV	296			
254 AISKAGAVPLQMEIGEGKDACVLEADADLDLVANIVYKGSYSQOKCTAVKVLIMESV	313			
297 ADELAETKSENVATLVSQDPFEDNATVTPVIDDNGADFESLIVVAPORGA---KE-----	348			
314 ADIVEKVKAKLALVTGPPPADSDITFVTVSSANFIEGLVMQKEGAFCEKKGATF	373			
349 INEFKRDRLTLTGLFDHVTLDMLKAMEEPPGPIPLITIRVADAEEAVALANKSPGLQSS	408			
374 CGEVRREBNLWPLBLDHVRDMMIAMEEPPGPVLPIVRINSVEGIIHHCANSFGLQGC	433			
409 VFTEDFOKAFDIANKLEVGTVIHNNKTGRGDFNFPFLGLKSGAGVOCIRYSIEMATNVK	468			
434 VFTKDINKAIIISDAMETGTVQINSAPARGDHPFQGLKDSGIGSGQITMSINMTVK	493			
469 STVLDM 474				
494 STVNL 499				

```

DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 508 AA; 54082 MW; FF027ACBF2743F58 CRC64;

Query Match
Best Local Similarity 47.6%; Score 1138.5; DB 2; Length 508;
Matches 231; Conservative 85; Mismatches 143; Indels 13; Gaps 5;

OY 11 GEMKSVN--QIEILSPIDSSLGFPVPMTRREEDHAMKAGREALPAMALTYERAQYL 68
DB 34 GMRVASAGSKVAIVNPTKRLTOYRQACTOGEVKNVMTATVAAQAMARTPLMKRAELL 93
OY 69 HKADIIERDKEIATVLAKEISKAVNASVTEVTPADLIRYAABEGIRLSTADEGGM 128
DB 94 HKAAAILKEHKTPRIAECLVKEIAKPAKDAISEVVRSGDLVSTAEGBVILDS---GXL 149
OY 129 ---DASTG--HKLAVIRROPVGIYLAIPNYPVNLSSSKTAPALIGENVMEFPTQG 182
DB 150 LVSDSPFGNERKKYCLSSKVPGLGVLAIPFPNVPVNLAVSKIGPALIAGNALVLKPTQG 209
OY 183 SVSGVLAKAFABAGLPAGVFNFTTGRGSEIGDYIYEHEENVFINTGSTPVGORIGKLA 242
DB 210 AVAALAHMWHCFILAGEPKKLINCVTGKSGEIGDFTLMHPGVNCISFTGG-DTGIAISKKA 268
OY 243 GMRPIMLEIGKDGAGIVLADADLDNAKOIVAGAVDYSQORCTAIRVLVVEADELAE 302
DB 269 GNVPLQMEIGKDGACVLEDDADLDVAANIVKGGFSGORCTAVKVLIMESVADIYE 328
OY 303 KISENVAKISVGDPPDNATVTPVIDNSADFTESLIVDAROKAKELNEFKDGLLTG 362
DB 329 KYAKAKLAKLTVPPEADSDITPVVTESSANFLBGLWDAKEKAGATPCQBYRREGLTWPL 388
OY 363 LFDHVLDMKLAMEEFPFGILPIIRVKDAEEAVALANKDFGLQSSVFRDPOKAFDIN 422
DB 389 LLDHVRPDMRIAMEEFPFGVLPVIRINSVEBGHHCNANPFLQCGVFTKDKINKALMISD 448
OY 423 KLEVGTVHINKTGRGPDNFPFLGLKSGAGVQGIKRSIEAMTNVKSIVLDM 474
DB 449 AMETGVQINSAPARGPDHFPFGQLKDSIGSGIGTNSIMMTKVKSTVINL 500

RESULT 27
O6RCS1 PRELIMINARY; PRT; 496 AA.
AC O6RCS1;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Non-phosphorylating GAPDH.
GN Name=ALDH1A5;
OS Physcomitrella patens (Moss).
OC Bryophyta; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC NCB1_Taxid=3218;
RN NCB1_Taxid=3218;
RP SEQUENCE FROM N.A.
RA Wood A.J., Reski R., Frank W.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY504666; AAS78753.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd. 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
SQ SEQUENCE 496 AA; 53347 MW; AEB641DC6C3BA6 CRC64;

Query Match
Best Local Similarity 46.8%; Score 1120.5; DB 2; Length 496;
Matches 229; Conservative 89; Mismatches 147; Indels 13; Gaps 5;

OY 5 YQNVNNGEKK--SSVQVQIILSPIDSSLGFPVPMTRREEDHAMKAGREALPAMALTYE 62
DB 16 FRYADGEWKKVSSGKSVGITNPSTLKQYKQACTODEVNVAVESAOAKIWAQTPIM 75

```

```

OY 63 ERAQYLHKAADIIEERDKEIATVLAKEISKAVNASVTEVTPADLIRYAABEGIRLSTSA 122
DB 76 KRAEALHFPALILKONKEIALVKEIAKHKALTEVNASGDLISABEGIRLILAE- 134
OY 123 DEGGM---DASTG--HKLAVIRROPVGIYLAIPNYPVNLSSSKTAPALIGENVMEF 176
DB 135 ---GKFLVSDSPFGNRRKYCLASKIPLGVLAIPFPNVPVNLAVSKIAPALIAGNAVVL 191
OY 177 KPTQGSVGLVLAFAFABAGLPAGVFNFTTGRGSEIGDYIYEHEENVFINTGSTPVQ 236
DB 192 KPTQGAASVLMWHCLHMAGPKGLVSAITGKSGEIGDFTLMHPGVNCISFTGG-DTGI 250
OY 237 RIGKLAGRPMLEIGKDGAGIVLADADLDNAKOIVAGAVDYSQORCTAIRVLVVEV 296
DB 251 AISRKAGVPLQMEIGKDGCCVLEDDADLELANNVIGKYSQORCTAVKVICVMSV 310
OY 297 ADELAEKISENVAKISVGDPPDNATVTPVIDNSADFTESLIVDAROKAKELNEFKDG 356
DB 311 ABEIVSKIVQMTKLTVMPEDNCDITPVVQSANFIOGLVEDAQAKAFHGEKREG 370
OY 357 RLTTGLDHTLDMKLAMEEFPFGILPIIRVKDAEEAVALANKSPGLQSSVFRDPOK 416
DB 371 NLIWPLLDVNTVPDMRIAMEEFPFGVLPVIRIKTYEBGHHHCNANPFLQCGVFTKDK 430
OY 417 AFDINKLEVGTVHINKTGRGPDNFPFLGLKSGAGVQGIKRSIEAMTNVKSIVLDM 474
DB 431 AILVSDAMESGITQINAPARGPDHFPFGQLKDSIGSGIGTNSIQMTTKKSTVINL 488

RESULT 28
O6MT99 PRELIMINARY; PRT; 471 AA.
ID O6MT99;
AC O6MT99;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (NADP) (EC 1.2.1.9) .
GN Name=gapdh; OrderedlocusNames=MSC_0509;
OS Mycoplasma mycoides (subsp. mycoides SC).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCB1_Taxid=44101;
RN NCB1_Taxid=44101;
RP SEQUENCE FROM N.A.
RA Westberg J., Persson A., Holmberg A., Goessmann A., Lundberg J.,
RA Johansson K.-B., Petersson B., Uhlen M.;
RT "The genome sequence of Mycoplasma mycoides subsp. mycoides SC type
RT strain PGIT, the causative agent of contagious bovine pleuropneumonia
RT (CBPP).";
RL Genome Res. 14:221-227(2004) .
DR EMBL; BX842643; CAE77137.1; -.
DR HSSP; P20000; IAG8.
DR GO; GO:0008886; F:glyceraldehyde-3-phosphate dehydrogenase (N. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd. 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 471 AA; 51976 MW; 5903AA20300770ED CRC64;

Query Match
Best Local Similarity 45.1%; Score 1079; DB 2; Length 471;
Matches 210; Conservative 108; Mismatches 150; Indels 2; Gaps 1;

OY 4 EYQNVNNGEKKSSVQVQIILSPIDSSLGFPVPMTRREEDHAMKAGREALPAMALTYE 63
DB 3 KFKALDDKLPDNNRILINVPDVSAGVSVLTQDINDAFTAKSSQKAMENTDLEK 62
OY 64 RAQYLHKAADIIEERDKEIATVLAKEISKAVNASVTEVTPADLIRYAABEGIRLSTAD 123

```

```
Db 63 RISILDKMKQILIDONKEELAQIIMSTFAKYKDCLEFVIRSVYIOTFPEVKNLTLLI 122
Qy 124 EGGRMDASTGHKLAVIRROPVGIVALIAPYVPVNLSGSKIALPALIGNVMMKPPTOGS 183
Db 123 DGAKEYAK--NKIGTFMRVAKGVGVAISPENYINLAVSKIFPCLVGNITVFKPATQGS 180
Qy 184 VSGVLVAKAPAEAGLPAGVFNITITGRGSEIGDIYVEHEVNFNFNGSTPVGQRIGKLAG 243
Db 181 LIGAKGELAYQNLKPGIFNVVITGRGREIGDDIITNKDLADFSIFGSEVVGRLLEIS 240
Qy 244 MRPIMLTEGGKDGIVLADADLDNAKOIVAGAVDYSQGRCTAIKRVLVVEEVADELAEK 303
Db 241 TKOVVLELGGKDPAIVLDDLDLEKAKETISGAFSGQRCCTAIKRVITTDKADQLVPL 300
Qy 304 ISENVAKLSVGDPEPDNAVTVPIVDNSADPIESLVVDARQKAKEINEFRDGRLLTPGL 363
Db 301 LKEKINKLTVGLPKDNCIDITPLIDOKTADPFVGLIDAKKKGAKIITIGKQENLITPPL 360
Qy 364 FDHVTLDMLKAMEEPFPIIPITRVDAEVAVALANKSDPGLQSSVFTRDFOKAFDIAK 423.
Db 361 VDHVTDMLKAMEEPFPIIPITRTSVQMIELANKSNFGLQASVYTKLDDALTIACK 420
Qy 424 LEVGVTHINNKTRGPDNPFPFLGKSGAGVQIGIRYSIEAMTVKSVL 473
Db 421 LEVGVTHINNKTRGPDNPFPFLGKSGAGVQIGIRYSIEAMTVKSVL 470
RESULT 29
Q6FLK7 PRELIMINARY; PRT; 472 AA.
AC Q6FLK7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, last annotation update)
DE NADP-dependent glyceroldehyde-3-phosphate dehydrogenase.
GN OrderedLocustNames=M1259;
OS Mesoplasma florum (Acholeplasma florum).
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Entomoplasmataceae; Mesoplasma.
OX NCBI_TaxID=2151;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H1 / ATCC 33453;
RA Birten B.W., Strange-Thomann N., Hafez N., Decaprio D., Fisher S.,
RA Butler J., Birkins T., Kodira C.D., Major J., Wang S., Nicol R.,
RA Nusbam C.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL:AE017263; F:oxidoreductase activity; IEA.
DR GO:GO:0016491; P:oxidoreductase activity; IEA.
DR GO:GO:0008152; P:metabolism; IEA.
DR InterPro:IPR002086; Aldehyd_dehydrog.
DR Pfam:PF00171; Aldehyd. 1.
DR PROSITE:PS00070; ALDEHYDE DEHYDR_CYS; UNKNOWN_1.
DR PROSITE:PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 472 AA; 51718 MW; 66EFA081AC4E8C04 CRC64;
Query Match 45.0%; Score 1077.5; DB 2; Length 472;
Best Local Similarity 44.8%; Pred. No. 2.1e-56;
Matches 211; Conservative 110; Mismatches 147; Indels 3; Gaps 2;
Qy 3 KEIQNVNGEMKSSVNOIEITSPIDSSLGFPVPMTRVEVDHMKAGREALPMAALTVY 62
Db 2 EKFKAIINNGEISNSGLIDMDPTDEVVAQVSAISQEIIDSAFKAKAAQKMEALIGIE 61
Qy 63 ERQVYIHKADIERDKKEIATYLAKEISKAYNAVTEVVRTDLIRYAEGRIST-8 121
Db 62 KRTEFLIRMDLLKKEEDLATMMHEIAYKDCLEVRRTAEYIDLITSEYNDLQVLT 121
Qy 122 ADEGKMADSTGHKLAVIRROPVGIVALIAPYVPVNLSGSKIALPALIGNVMMKPPPTO 181
Db 122 FDGNSK--GVTEIDVIAEYKRIAGVGVGISPPYIPINLAVSKLAPGLITNTVFEPKAT 179
```

```
Qy 182 GSVGVLVAKAPAEAGLPAGVFNITITGRGSEIGDIYVEHEVNFNFNGSTPVGQRIGKL 241
Db 180 GSVGVLVAKAPAEAGLPAGVFNITITGRGSEIGDIYVEHEVNFNFNGSTPVGQRIGKL 239
Qy 242 AGMRPIMLTEGGKDGIVLADADLDNAKOIVAGAVDYSQGRCTAIKRVLVVEEVADELA 301
Db 240 SSSKDLVLELGGKDAIILDEHNLENLADYAGAFSGQRCCTAIKRVITTDNDADKT 299
Qy 302 EKISENVAKLSVGDPEPDNAVTVPIVDNSADPIESLVVDARQKAKEINEFRDGRLLTP 361
Db 300 PLIKAEVAKLTVGLPENENIITPMIDKKSADPVTDLINDAVKGAFTLVGSTRDKNLQP 359
Qy 362 GFDPHTLDMLKAMEEPFPIIPITRVDAEVAVALANKSDPGLQSSVFTRDFOKAFDIA 421
Db 360 TLLDVTVMYNAWMEPFPFPIIPITRINIEIMIIVANKSNFGLQISYISKVDLAVKVA 419
Qy 422 NLEVGVTHINNKTRGPDNPFPFLGKSGAGVQIGIRYSIEAMTVKSVL 472
Db 420 EQLEGVTHINNKTRGPDNPFPFLGKSGAGVQIGIRYSIEAMTVKSVL 470
RESULT 30
Q8EVT9 PRELIMINARY; PRT; 478 AA.
AC Q8EVT9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE NADP-dependent glyceroldehyde-3-phosphate dehydrogenase.
GN OrderedLocustNames=MYPB4710;
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22354719; PubMed=12466555; DOI=10.1093/nar/gk1667;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5283-5300(2002).
DR EMBL:AP004172; BAC44260.1; -.
DR HSSP; Q59931; 1EUDH.
DR GO:GO:0016491; P:oxidoreductase activity; IEA.
DR GO:GO:0008152; P:metabolism; IEA.
DR InterPro:IPR002086; Aldehyd_dehydrog.
DR InterPro:IPR002332; PII_glnb_UMP_S.
DR Pfam:PF00171; Aldehyd. 1.
DR PROSITE:PS00070; ALDEHYDE DEHYDR_CYS; UNKNOWN_1.
DR PROSITE:PS00687; ALDEHYDE DEHYDR_GLU; UNKNOWN_1.
DR PROSITE:PS00496; PII_glnb_UMP; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 478 AA; 52311 MW; ADDA6110273087C5 CRC64;
Query Match 43.8%; Score 1047.5; DB 2; Length 478;
Best Local Similarity 46.5%; Pred. No. 1.3e-54;
Matches 218; Conservative 90; Mismatches 152; Indels 9; Gaps 3;
Qy 8 YNNGEMKSSVNOIEITSPIDSSLGFPVPMTRVEVDHMKAGREALPMAALTVYRAQY 67
Db 11 YNNGKRVNSEKIPVISTIDGSLIGVAAISKQIDIMFEGAHLEFNLNRNLTPKWRIRK 70
Qy 66 LKRAADIIRDEKEIATYLAKEISKAYNAVTEVVRT----ADLIRYAEGRISTSD 123
Db 71 IKEPAEYFAEKEPFLTLMSYELGKSYNDALKEVERSYNIFETLVEYKEKEVNSVIGP 130
Qy 124 EGGRMDASTGHKLAVIRROPVGIVALIAPYVPVNLSGSKIALPALIGNVMMKPPTOGS 183
Db 131 EVNKKIKGKTGYVYV---PVGVLVLAIAFPYPINLGLAKIIPILLVGNITVFKPATQGS 186
Qy 184 VSGVLVAKAPAEAGLPAGVFNITITGRGSEIGDIYVEHEVNFNFNGSTPVGQRIGKLAG 243
```


Db 187 LVSSOLAKYFDQANFAGVFNLTGKSGEIGDIILENKRIQATFTGSSDIFKLASKLP 246
QY 244 MBPIMLEIGCGDAGIVLADADLDNAAKOIVAGAYDSGQRCTAIRKRVLVVEADLEAK 303
Db 247 MBPVLLEIGGKDAALITNNADVELAKEITIKAFGYSQRCTAIRKRVLTKEIADLLIH 306
QY 304 ISENAKLSVGDPPDNATVPVIDNSADFIESLVADARQAKELNEFKRG-RLTFPG 362
Db 307 LIKEAQKLVGNPILNPPDITPLIDRAVDNFNIDLDDALNHGANLVYGINIGYNILRHT 366
QY 363 LFDVHTLDMKLAMEBPFGLPIIRVKAEEVAVALANKSDPGLQSSVFTFRDPOKADIAN 422
Db 367 ILDNVSTSKLAMEBPFGLPIIRVSTINEAIAINASEYGLQGSVFTFKDLEARTIAK 426
QY 423 KLEVGTVHNNKTGRGPDNPFPLGLKSGAGVQGRYSIEAMTNYSIV 471
Db 427 YLDGTVINIRGSSRGPDIFFPIGLKSGVQGRISDKAMTRVKGIV 475

RESULT 31

Q8LK61 PRELIMINARY; PRT; 496 AA.
AC Q8LK61
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=22275840; PubMed=12387887; DOI=10.1016/S0014-5793(02)03455-5;
RA Bustos D.M., Iglesias A.A.;
RT "Non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase is post-translationally phosphorylated in heterotrophic cells of wheat (Triticum aestivum).";
RL FEBS Lett. 530:169-173(2002).
DR EMBL, AF521191; NAMB7679.1; -.
DR HSSP; Q59931; IEUH.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 496 AA; 52872 MW; 20F9D07F1B840FB CRC64;

Query Match 42.9%; Score 1027.5; DB 2; Length 496;
Best Local Similarity 45.0%; Pred. No. 2.5e-53;
Matches 215; Conservative 78; Mismatches 172; Indels 13; Gaps 4;

QY 5 YQNVNGEKKSVN--QIILISPIDSSIGFVPAMTRREVDHAMKAGBALPAMALTVY 62
Db 16 YKYVADGEWRASASGKVAIVNPTTRQTYRVOACTQEEVNKAKMDAKVAKXKWARTPLM 75
QY 63 ERAQVILHKAADIIEKDKEIATVLAKEISKAVNASVTEVTRADILIRAAEGIRLSTSA 122
Db 76 KRAELIHLKAAALIKHEKTPIAECLVKEIAKPAKDAVSEVVRSDLVSTYAESEVRLIG- 134
QY 123 DEGGKMDASTG-----HKLAVIRQPGIVLAIAPYVNVLSGSKIAPALIGNVNF 176
Db 135 ---GKLXVNSBSPGNERNKYCLSSKIPLGVALIAPFPYIPANLASKIGPALIAGNALVL 191
QY 177 KPTPGSVSGVLAKAFAGAGLPAGVFNTITGRGSGIDGYVHEHEVNFINTGSPVQO 236
Db 192 KPTQGAVALHMHVCHFLAGFPKGLISCVTGKSGEIGDPLTMHPGVNCISFTG-DTGI 250

QY 237 RIGKLGNRPIMLEIGCGDAGIVLADADLDNAAKOIVAGAYDSGQRCTAIRKRVLVVEV 296
Db 251 AISKAGAVPLMXMLXGDACTVXEXADLDLVXANIVVGGRFSYSGQRCAVAVVILIMEXX 310
QY 297 ADELAEKISENAKLSVGDPPDNATVPVIDNSADFIESLVADARQAKELNEFKRG 356
Db 311 ADVXVZKVNAKLAKIKVPPPEXXXDITVVTSSANFVXGGBXCKGRXNFXLPVQDEK 370
QY 357 RLTFPGLDVHTLDMKLAMEBPFGLPIIRVKAEEVAVALANKSDPGLQSSVFTFRDPOK 416
Db 371 XLIMPDLHDVHRPDXIXAMKEBFGFVLVIRINSVEGIIHCNANSFGLQGVFTFRDINK 430
QY 417 AFDIANKLEVGTVINNNKTGRGPDNPFPLGLKSGAGVQGRYSIEAMTNYSIVDM 474
Db 431 AIMISDAMESGTVOQNSAPARGPDHFPQGLKDGSGIGGIGTNSINMTKYSIVINL 488

RESULT 32

Q8LSJ9 PRELIMINARY; PRT; 496 AA.
AC Q8LSJ9
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Endosperm;
RX MEDLINE=22275840; PubMed=12387887; DOI=10.1016/S0014-5793(02)03455-5;
RA Bustos D.M., Iglesias A.A.;
RT "Non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase is post-translationally phosphorylated in heterotrophic cells of wheat (Triticum aestivum).";
RL FEBS Lett. 530:169-173(2002).
DR EMBL, AF521190; NAMB7678.1; -.
DR HSSP; Q59931; IEUH.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd; 1.
KW Oxidoreductase.
SQ SEQUENCE 496 AA; 52869 MW; 8C1B1F15FF33881B CRC64;

Query Match 42.9%; Score 1026.5; DB 2; Length 496;
Best Local Similarity 45.0%; Pred. No. 2.5e-53;
Matches 215; Conservative 78; Mismatches 172; Indels 13; Gaps 4;

QY 5 YQNVNGEKKSVN--QIILISPIDSSIGFVPAMTRREVDHAMKAGBALPAMALTVY 62
Db 16 YKYVADGEWRASASGKVAIVNPTTRQTYRVOACTQEEVNKAKMDAKVAKXKWARTPLM 75
QY 63 ERAQVILHKAADIIEKDKEIATVLAKEISKAVNASVTEVTRADILIRAAEGIRLSTSA 122
Db 76 KRAELIHLKAAALIKHEKTPIAECLVKEIAKPAKDAVSEVVRSDLVSTYAESEVRLIG- 134
QY 123 DEGGKMDASTG-----HKLAVIRQPGIVLAIAPYVNVLSGSKIAPALIGNVNF 176
Db 135 ---GKLXVNSBSPGNERNKYCLSSKIPLGVALIAPFPYIPANLASKIGPALIAGNALVL 191
QY 177 KPTPGSVSGVLAKAFAGAGLPAGVFNTITGRGSGIDGYVHEHEVNFINTGSPVQO 236
Db 192 KPTQGAVALHMHVCHFLAGFPKGLISCVTGKSGEIGDPLTMHPGVNCISFTG-DTGI 250
QY 237 RIGKLGNRPIMLEIGCGDAGIVLADADLDNAAKOIVAGAYDSGQRCTAIRKRVLVVEV 296
Db 251 AISKAGAVPLMXMLXGDACTVXEXADLDLVXANIVVGGRFSYSGQRCAVAVVILIMEXX 310

[illegible]

RESULT	33
07NBXS	
ID	07NBXS
AC	PRELIMINARY;
DT	01-MAR-2004 (TREMBlrel. 26, Created)
DT	01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE	PuLA.
GN	ORFNames=MGA_0860;
OS	Mycoplasma gallisepticum.
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX	NCBI_TaxId=2096;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=R;
RX	MEDLINE=22830409; PubMed=12949158; DOI=10.1099/mic.0.26427-0;
RA	Papazisi L., Gorton T.S., Kutish G., Markham P.P., Browning G.F.,
RA	Nguyen D.K., Swartzell S., Madan A., Mahitras G., Geary S.J.;
RT	"The complete genome sequence of the avian pathogen Mycoplasma
RT	gallisepticum strain R(low)";
RL	Microbiology 149:2307-2316(2003).
DR	EMBL; AE016967; AAP56485.1; -.
DR	HSSP; Q59931; IEUH.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0008152; P:metabolism; IEA.
DR	InterPro; IPR002086; Aldehyde_dehydrog.
DR	Pfam; PF001071; Aldehyd. 1.
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.
FW	Complete proteome.
QO	SEQUENCE 525 AA; 58881 MW; 9DA7CF18215BA981 CRC64;

Query Match	39.4%;	Score 944;	DB 2;	Length 525;
Best Local Similarity	45.0%;	Pred. NO. 2.3e-48;		
Matches 210;	Conservative 76;	Mismatches 173;	Indels 8;	Gaps 5

```

Qy 5 UNDEBMSXVNOQIILISPIDDSSLPFPMANTREVDHAKAGREALPANA-ALTYYEAOY 67
Db 60 INNOVEGSSKLIEIKSPIDHEISGSPYGAETIEIDAEKADIAKSSKSGYERAKI 119
68 LHKAAIDIERDEKEIATVLAKESKAYNA5VEEVRTDLIR---YAAE6GIRLSTADE 124
120 L-KRAGQLDHAHKEEPANILTDNIYAKYNSLAEVTRSIQYIHDTISYEEMITHPLWDE 178
Qy 125 GGRNDASTGHKIAVIRQPIGYILAIAPNTYPVNISGSKIAPALIGCNVNMFKPPTGGSV 184
Db 179 --KIH-RVKEKVKFVREPLGVALNISPANYPLNPLPSMIETLLAGTIVVYKATQDAI 235
185 SGVLAKAFAPAEAGLPAGVEFNTITGRGSEIIGDYIVHEEYVNFINTGSPYQOIRGYKLAM 244
236 IGRILAHLPFOAGFPDGOVQCVGGLGREIGDKLNTNNKHKSISFGSPFVGLSLLKQSAV 299
Qy 245 RPIWLEIIEGSDAGIYIADADLDNAALQIYAGAYVDVSGORCTAIRKIVYVEEABELEAKI 304
Db 296 GNISLIEIGSDAALVYSDVQKRTTIEIKGAYSSGQCTAIRKIVFVKXNDADBEVQOL 355
Qy 305 SENYAKLSVDPDPDMATVTPVVIDNDSADRIE5LVDARQGAKEINERKDGRLITGQLF 364
356 KKEVEKIEHLNPNFNPBSLVVIDIOHANYIKELFPDAVTKRAIVYTGDFPKNCVYAILL 415

```

QY 365 DHTLDMKAMEEPFPILPIPIIRVDAAEVAIVANKSPFGQSSVFTRDQKAFDLANKL 444
 Db 416 DHATTMMRAAMEEPFPILPIPIIRVNEAMEWLNHNSQSYGAQSIPIFSDQKDELANOL 475
 QY 425 EVGTVAHNKTKGRGPDNPFPLGLKSGAGVGQIRYSIEAMNNVKSIV 471
 Db 476 EAQTINNNRSSRGPEFPFPMGVKDSGVGVGQIRDTLTISVRYRGFV 522

RESULT	34		
Q9PQC9		PRELIMINARY;	PRT; 475 AA.
AC	Q9PQC9;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)		
DE	MADP-dependent glyceraledehyde-3-phosphate dehydrogenase.		
GN	Name=gapb; OrderedAccession=U0362;		
OS	Ureaplasma parvum (Ureaplasma urealyticum biotype 1).		
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.		
OX	NCBI_TaxID=134821;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 700970;		
RA	MEDLINE=20500219; PubMed=11048724; DOI=10.1038/35037619;		
RX	Glase J.1., Leftkowitz E.J., Glase J.S., Heiner C.R., Chen E.Y.,		
RA	Cassell G.H.;		
RT	"The complete sequence of the mucosal pathogen Ureaplasma		
RT	urealyticum."		
RL	Nature 407:757-762(2000).		
DR	EMBL; AE002133; AAF30771.1; -.		
DR	HSSP; Q59931; 1EUH.		
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.		
DR	GO; GO:0008152; P:metabolism; IEA.		
DR	InterPro; IPR002086; Aldehyd_dehydrog.		
DR	Pfam; PF00171; Aldehyd_1		
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.		
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.		
FW	Complete proteome.		
Q0	SEQUENCE 475 AA; DDB00D06576603752 CRC64;		

Query Match	36.3%	Score 868.5;	DB 2;	Length 475;
Best Local Similarity	39.7%	Pred. No. 6.8e-44;		
Matches 190; Conservative	95;	Mismatches 177;	Indels 17;	Gaps 5

```

Oy 3 YONVYVNGEKKSVNOJELISPIDDSSLCGFVPMATRE -EVDHAMAGGREALPMAALTYVE 63
Db 3 YKLLINGAVDVAKEKLPVYNPSNNQI1AYVYNHNEBENINTIFENAHIAALKRPDPICKY 62
Oy 64 RAQYLHKADDI1ERDBKEIATVLAJESKAVNASVTEVTRADILIRAAEGR1R1STA- 122
Db 63 RCDLLK1KADDKDENHQBOLAQI1TSEIKGLKDSLIEVERADY1RETVEYOKLMOKPI 122
Oy 123 --DEGKMDSATGCH---KLAVTRBPQGIYLA1APYVYNLSSGSIAPALIGSNVWF 176
Db 123 I1PB-----IVHHVKKVKA1TYK11P1GVVLA1C1P1NY1N1L1SLA1PALV1SGNS1VY 175
Oy 177 KPPTGQSVSGVLVLAFAEAGHPAGVFN1T1GRGSE1GDIY1VEHEBVNF1NFTGSP1VGO 236
Db 176 KPPTQSG1L1GIR1SEL1VHEVGP6KVNVC1L1TEAR1TGDD1LV1TKYKAL1S1FTG6P1KVN 235
Oy 237 R1G1K1LAKMP1M1LE1G1G1D1AG1V1LAD1AD1DMA1Q1YA1G1ADV1G1G1C1T1AK1V1LVEBV 236
Db 236 H1A1ET1SK1SLV1E1TG1G1D1PALV1DD1AD1F1EA1N1E1Y1G1A1G1F1G1G1C1T1AK1R1F1V1SHKN 235
Oy 297 ADELAEK1SENVAK1SVG1D1P1N1ATV1T1VID1D1N1S1AD1F1ES1LV1DAR1OKGA--KE1NE1FKR 354
Db 296 HD1LV1NL1K1K1D1AL1TV1GL1PQ1N1P1T1T1L1N1S1N1K1L1S1VEB1A1IKK1A1YHOK1LYNE 355
Oy 355 DGR1LL1P1G1L1FDH1V1L1DM1K1LAE1E1P1P1L1P1I1RYK1DA1E1A1V1A1N1K1S1D1F1G1L1S1V1F1TRD1F 414
Db 356 KN1L1H1P1L1V1D1V1T1K1EM1R1VA1E1E1P1G1P1L1P1T1N1S1O1E1A1D1L1N1O1Y1G1A1C1IF1T1YV 415

```



```

DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Aldehyde dehydrogenase.
GN OrderedLocustNames=PAE2480;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
NCBI_TaxId=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RA MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fltz-Gibson S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum."
RL EMBL; AEO09879; AL64225.1; -.
DR HSSP; Q59931; 1EUN.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR Pfam; PF00171; Aldehdh; 1.
KW Complete proteome.
SQ SEQUENCE 478 AA; 52861 MW; 9FC9409116223E7 CRC64;

Query Match 32.8%; Score 786; DB 2; Length 478;
Best Local Similarity 37.1%; Pred. No. 5,9e-39;
Matches 176; Conservative 96; Mismatch 192; Indels 10; Gaps 4;

QY 6 QNYVNGEMKSNVQVIEILSPIDSSIGFVPMTRREEDHMKAGREALPMALITYERA 65
DB 5 KNINFEFIEVERRRPHSHSIDSVIAETVSKREDKAAIDAYDLKMSQULPIKRA 64
QY 66 QYLHAKADIIERDKKEIATYLAKEISKYNAVTEVTRADLIRYAEBCIRLSTADEG 125
DB 65 EYUUKLEYIKSRBEELINILMVEGGIYKKWGEVFTLRVNRNAB-----LARHYG 118
QY 126 GK-MDASTGKHLAVIRROPYGLIALAPYVPVNLSSGSKTAPLIGNVMPFPPOGVS 184
DB 119 GKTLQSPAEQVSVMPFRKPGVGVITPMVYPLISIKKKTAHMLATGNTIIVYRPASSTPV 178
QY 185 SGLVLKAFEAAGLPGAVENTITGRGSEIGDYVEHEVNFIFNGTSPVQRIKLAG- 243
DB 179 TGRKIAMHQAQFPKQVNLVVGPGSTVGEIVVNNKQSHVFTGESATGRILASRAGS 238
QY 244 -MNPIMELGKGDAGIVLADADVNNAKQIVAGAYDYSQRCYAIKRVLVEEVADELA 302
DB 239 ALKTVLELGGSDPLIILDDADNLAVRIVFGAFPHQGICTSSKRILVHERVYDVFK 298
QY 303 KISEENAKLSVGPFPDNATVT-VIDDNSADPESLVVNDAROKAKELEFRKDRLLTP 361
DB 299 KPFERVSQLEKVGPRPRRTVEQGLISPRQADEMERFYQDVVSQGLALTGSKRAGYFWP 358
QY 362 GLFDHTLDMKLMEEPFGPIPLIRYKDAEVAVALANKSDPLQSSVFTDRQKAFDA 421
DB 359 TVVVDVDRNRIRMEEPFIRVVPVKNDEVAEVAANDVEGASAAVTTNINRAFKKA 418
QY 422 NKLVEGTVINNTKGRPNPFLGLKSGAGVQGIIRYSIEAMTNVYSIVLDMK 475
DB 419 EALIESGVHINDVTMLAESHVPPGGIKASFGREGGWSVHETTYRMLITTLR 472

RESULT 38
DHAI_BACSU STANDARD; PRT; 488 AA.
AC P42236;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Probable aldehyde dehydrogenase ycdB (EC 1.2.1.3).
GN Name=ycdB; OrderedLocustNames=BST02470;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxId=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA MEDLINE=95219079; PubMed=7704254;
RA Ogawa K.-I., Akagawa E., Nakamura K., Yamane K.;
RT "determination of a 21548 bp nucleotide sequence around the 24 degrees
RT region of the Bacillus subtilis chromosome."
RL Microbiology 141:269-273(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunze F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Brun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruch C.V., Caldwell B., Capuano N.J., Carter N.M.,
RA Chou S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denton F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaeser P., Goffeau A., Gollighly E.J., Grandi G.,
RA Guiseppl G., Guy B.J., Haga K., Hatech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaeffer-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.C.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moesli D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Roche E., Roche B.,
RA Rose M., Sadate Y., Sato T., Scanlan E., Schleich S., Schroeder R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serro P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Teipstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassaretti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Wellensgger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasunoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumberstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -1- CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
CC -1- PATHWAY: Ethanol utilization; second step.
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D30808; BAA06468.1; -.
CC EMBL; Z99105; CAB12041.1; -.
CC PIR; G69752; G69752.
CC HSSP; P51977; 1BXS.
CC Subtilisin; Bg1159; ycdB.
CC InterPro; IPR002086; Aldehyde_dehyd.
CC Pfam; PF00171; Aldehdh; 1.
CC TIGRPFAM; TIGR01804; BMSDH; 1.
CC TIGRPFAM; TIGR01722; BMSDH; 1.
CC TIGRPFAM; TIGR01780; SSADH; 1.
CC PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
CC PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
CC Complete proteome; Hypothetical protein; NAD; Oxidoreductase.
KW NP BIND 233 NAD (ADP part) (By similarity).
FT ACT_SITE 255 By similarity.
FT ACT_SITE 289 By similarity.

```

```

SQ SEQUENCE 488 AA; 52414 MW; 9C37B32EF51D59B2 CRC64;
Query Match 31.9%; Score 763; DB 1; Length 488;
Best Local Similarity 36.9%; Pred. No. 1.4e-37;
Matches 170; Conservative 94; Mismatches 179; Indels 18; Gaps 6;

QY 5 YQNYNGEN--KSSVNOIILSPID--DSSIGFVPANTREEDHAMKAGREALPMAALTV 61
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 10 YNFMINGEWMKSSQSGDMVKENDPADVNDIVGQNSTADVERAVALAANEAKTAMKX 69
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 YERAQVILHKAADIIERDKERIATVLAKELISKAYNASVTEVPTADILIRAAEGRILSTS 121
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 70 AARGQVLYTKADIMEERLEIAACATRENGKTLPEAKGETARIGATILRYAAGGKEXT-- 127
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 ADEGGMDASTGKHLAVIRROPVGIATAPVNYPNVLSGSKIPALIGANVMPFPPTQ 181
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 128 ---GDVPESTDKDALMFTTRVPLGVGVISPMNFPALIRPKAPALVIGNVVKEPAT 184
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 GSVGLVIAKAPAEAGLPAGVFNITIGRGEISIDYIYEHEBVNFINTGSTPVGQRIGKL 241
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 185 TAVTCAKIIACEFEAGLPAGVINLVTPGPGSVGQGLAEHDGVAAVFTSGNQKILIGGA 244
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 AGMR--PIMLEIGCKRAGIVLADADIDNAKQIVAGAYVSGQRCTAIRKVLVEEVADE 299
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 245 ALARGAKYQLEWGGKNPVIVADADLEAAAEAVITGAFRSTGQCKTATSRVIVQSGIYER 304
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 LAEKISENVAKLSVGDPF--DNATVTPVIDNSADPIESLVLVDAROKAKEL-----NE 351
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 305 FPEKLLQRTKDTITIGSLDEKEDVWMPVASKNQDNCGLSTIEKKQKQGAULLIGEGLENG 364
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 352 FPRDGLLTPGLFDHVTLLDKLAMEEPFGPIILPIIRVKDAEBAVALANKSDPGLQSSVPT 411
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 365 KYONGYVOPALFEDNTSEMTIAQOEIRFGVIALITVDSIEBALNIAIVDKGLASISITP 424
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 412 RPOQKAFDIANKLAEVCTVTHINKTGKGPDPNFPPLGLKSGGA 452
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 425 ENIGRLSFIIDEIDAGLVIRINAEAGVELQAPPGKMKQSSS 465
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 39
YELL_METUA STANDARD; PRT; 463 AA.
ID YELL_METUA STANDARD; PRT; 463 AA.
AC 058806;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical aldehyde-dehydrogenase like protein MJ1411 (EC 1.2.1.-).
GN OrderedlocusNames=MJ1411.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxId=2190;

SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geochagen N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Utecherback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RT J. Bacteriol. 187:2128-2138 (1995).
RT -i- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; U67581; AA899418.1; -.
DR PIR; B64476; B64476.
DR HSSP; Q59931; 1E0H.
DR TIGR; MJ1411; -.
DR Interpro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00172; Aldehyd, 1.
DR TIGRFAMs; TIGR01804; BADH, 1.
DR TIGRFAMs; TIGR01780; SSADH, 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; FALSE NEG.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; FALSE NEG.
KW Complete proteome; Hypothetical protein; Oxidoreductase.
FT ACT_SITE 240 By similarity.
FT ACT_SITE 274 By similarity.
SQ SEQUENCE 463 AA; 51135 MW; 46F26568595B6A5 CRC64;

Query Match 31.7%; Score 759.5; DB 1; Length 463;
Best Local Similarity 37.1%; Pred. No. 2.2e-37;
Matches 175; Conservative 95; Mismatches 183; Indels 19; Gaps 8;

QY 8 YNNGEWSKSSVNO--LEILSPIDSSLGFPVAMTREEDHAMKAGREALPMAALTVYERA 65
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2 FIDGKQ---INREDMDVINPYSLIEYIKKI PALSREAEKALIDTAEKYKEMKMLPTTKRY 58
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 66 QYLHKAADIIERDKERIATVLAKELISKAYNASVTEVPTADILIRAAEGRILSTSADGG 125
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 59 NILNIAIAQIKERKEELKILATIDAGKPIKQARVEVERISGTFKLA--FYVEKHREVV 115
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 126 GKMDASTGKHLAVIRROPVGIATAPVNYPNVLSGSKIPALIGANVMPKPTQGSVS 185
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 116 IPSD----DRLIFTRREPVGIGATPPNFPPLNSAHKILAIATANGVIVHHPSSKAPLV 171
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 186 GLVIATK---APAEAGLPAGVFNITIGRGEISIDYIYEHEBVNFINTGSTPVGQRIGKL 241
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 172 CIELAKIENALKKKNVPLGVNLLTGAEVGEIVNEKVMNLSFGSSKVGELITKK 231
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 AGMRPIMLEIGCKRAGIVLADADIDNAKQIVAGAYVSGQRCTAIRKVLVEEVADELA 301
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 232 AGFKKIALELGGVNNIVLKADLNKAVNALIKSFYAGOVCSVGMILVDESIADFT 291
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 302 EKISENVAKLSVGDPFDNAT--VTPIYDINSADPIESLVLVDAROKAKELNFKEDGRLLT 360
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 232 EMFVVKAVLANVGNLDEKTDVGPLISVEHAEWKEVVKAIIDEGKLLDGGKRKALFY 351
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 361 PGLFDHVTLLDKLAMEEPFGPIILPIIRVKDAEBAVALANKSDPGLQSSVPTRDPOKAFDI 420
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 352 PFIIE--VDRDNLCTETFPAPIPIRTNE--EEMIDIANSLEYGLHSAIFINDIKSLKF 409
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 421 ANKLEVGIVTHINKTGKGPDPNFPPLGLKSGAGVQGIKYSIEAMTNVSYLV 472
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 410 AENLEFGVINDSSLFRODNNPFGVKKSGLGREGVAYAMEWSNIXITII 461
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 40
PRELIMINARY; PRT; 470 AA.
ID Q97XA5 PRELIMINARY; PRT; 470 AA.
AC Q97XA5;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase, NADP dependent (GapN-2) (EC
DE 1.2.1.9).
GN Name=GapN-2; OrderedlocusNames=SSO1842;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxId=2287;
RN (1)
SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;

```

RC MEDLINE=21332296; PubMed=11427726; DOI=10.1073/pnas.141222098;
RX She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moor A., Etraugo G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.,
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.",
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL: A8006793; AAK42037.1; -.
DR PIR: P90347; P90347.
DR HSSP: O57693; 1KY8.
DR GO: GO:0016491; P:oxidoreductase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002086; Aldehyde_dehydrog.
DR Pfam: PF00171; Aldehyd. 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 470 AA; 51603 MW; 6C22E206D5B8326C CRC64;

Query Match 31.4%; Score 752; DB 2; Length 470;
Best Local Similarity 36.5%; Pred. No. 6.2e-37;
Matches 172; Conservative 96; Mismatches 187; Indels 16; Gaps 7;

QY 16 SYNQIITLSPIDSSLSGFVPAMTRREVDHAKAGREALPAMALTYERAQYLHKAADII 75
DB 3 SLEVEIRSPNKLKVGIVKMSKDEVRGEIEBAYGFEIISMPYKRTALIRKSEIL 62
QY 76 EHDKEIATVLAKEISKAVNASVTEVVRTADILRYAAEE-GIRLSTSDGCKMD- 130
DB 63 EKKQERLALAMEAKRPIKDSRVEVLRASRLFRQAEEVGLVLE--GNTRVDAYEYP 119
QY 131 -SFGHLAVIRROPVGIYLAAPYNYPNVLSGSKIAPALIGNVNMEKPEPTQSGVGLV 189
DB 120 PGNENRIVLSTREPIGVTAALPFPNPINSFAHKVAPALAVGSSVYVKSINPLAAYM 179
QY 190 AARFAPAGIPAGVFNTITGSGEIGDIYVEHEVNFINTGSTPVGQRIKGLA--GMRP 246
DB 180 KILVAGAGIPDSAVRVVGTSGSEIGDEIITHPLVGLITLTGSTGTGLAKSKAISLGR- 238
QY 247 IMLELGGKAGIYLAADLDMNAKQIYAGAYDSGCGCTAIKRVLYVEEVADELAEKISE 306
DB 239 IIMELGSSPILILEPANTERRASSIARARFEYAGQNCNAGKRIIVROEYVDYFVAFNE 298
QY 307 NNAKLSVGDPEFNAT-VTPVIDNSADFIESTLVVDAROKAKE--LNEFKRDGRLTTPGL 363
DB 299 KAKALKVGEPLDETTIVGVPIVINKESVENINSVLEDAKVGGRVAILNRGPESGFPFLTM 358
QY 364 PDHVTLLDMKLAMEEPFGPIIPITRVKDAEBAVAIANKSDFGLOSVFTRDFOKAFDIANK 423
DB 359 VTNPSIDMLVLTKEVGFPIVPIVSVKSDBEAIRIANSTVEGLQSAIFTDVNRALKLSRE 418
QY 424 LEVGYTHINKTGRGPNFPFLGLKSGGAGVQIRISIEAMTVKSIIVLDM 474
DB 419 LKFGAVIINDSTRLRWDSLPEFGFKTKTIGREGVRETMLMTEKNLIAITL 469

RESULT 41
O976X5
ID O976X5 PRELIMINARY; PRT; 468 AA.
AC O976X5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE 468aa long hypothetical aldehyde dehydrogenase.
GN OrderedLocustNames=ST0064;
OS Sulfolobus tokodaii.
OC Archaeae; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCB1_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kusuda N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanggiti M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.,
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.",
RL DNA Res. 8:123-140(2001).
DR EMBL: APO00981; BAB65021.1; -.
DR HSSP: O57693; 1KY8.
DR GO: GO:0016491; P:oxidoreductase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002086; Aldehyde_dehydrog.
DR Pfam: PF00171; Aldehyd. 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 468 AA; 51257 MW; 523C477E7F0F3BE CRC64;

Query Match 31.4%; Score 751; DB 2; Length 468;
Best Local Similarity 36.7%; Pred. No. 7.1e-37;
Matches 171; Conservative 90; Mismatches 191; Indels 14; Gaps 6;

QY 20 IITLSPIDSSLSGFVPAMTRREVDHAKAGREALPAMALTYERAQYLHKAADITERDK 79
DB 5 IEKSPNKLKVGIVKMSKDEVRGEIEBAYGFEIISMPYKRTALIRKSEILREQ 64
QY 80 EELIATVLAKEISKAVNASVTEVVRTADILRYAAEEGIRLSTSDGCKMD- 134
DB 65 EELIARLAMEAKRPIKDSRVEVLRASRLFRQAEEVAIVLEG--KNYRVDAEYPPGNE 122
QY 135 KLAIVIRROPVGIYLAAPYNYPNVLSGSKIAPALIGNVNMEKPEPTQSGVGLVAKAPA 194
DB 123 RIVISTREPIGVTAALPFPNPINSFAHKVAPALAVGSSVYVKSINPLAAYM 182
QY 195 EAGLPAGVFNTITGSGEIGDIYVEHEVNFINTGSTPVGQRIKGLA--GMRPIMLE 251
DB 183 EAGLPDSAVRVVGTSGSEIGDEIITHPLVGLITLTGSTGTGLAKSKAISLGR- 241
QY 252 GKKDAGIYLAADLDMNAKQIYAGAYDSGCGCTAIKRVLYVEEVADELAEKISEVNT 311
DB 242 GGSDDPIVLEDNIDRASSIARARFEYAGQNCNAGKRIIVREIYDFKVKAFKKVKAL 301
QY 312 SVGDPEFNAT-VTPVIDNSADFIESTLVVDAROKG--AKEINFRDGRLLTTPGLFDHT 368
DB 302 KKGDPDESDTIDGPIVINGESVKKALKLEDAQSKGNVAVLNKGPETGFFFLSLVTNS 361
QY 369 LDMKLAMEEPFGPIIPITRVKDAEBAVAIANKSDFGLOSVFTRDFOKAFDIANKLEVT 428
DB 362 LDMVLKTEIFPIPIVPIVSVKSDBEAIRIANSTVEGLQSAIFSDVNRALKIAKELKFGA 421
QY 429 VHNKKTGRGPNFPFLGLKSGGAGVQIRISIEAMTVKSIIVLDM 474
DB 422 IINDSTRLRWDSLPEFGFKTKTIGREGVRETMLMTEKNLIAITL 467

RESULT 42
O8EMK2
ID O8EMK2 PRELIMINARY; PRT; 486 AA.
AC O8EMK2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Aldehyde dehydrogenase (EC 1.2.-.-).
GN OrderedLocustNames=O82840;
OS Oceanobacillus theyensis.
OC Oceanobacillus; Bacteriia; Firmicutes; Bacillales; Bacillaceae; Oceanobacillaceae;
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillaceae;
OX NCB1_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTB831;

RA MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gke526;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of *Oceanobacillus theyensis* isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments";
RT Nucleic Acids Res. 30:3927-3935(2002).
RL EMBL; AF004602; BAC14796.1; -.
DR HSP; P51977; 1BX5.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde_dehydrog.
DR Pfam; PF00171; Aldehyd_1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 486 AA; 52280 MW; 39A5A1941B4013B CRC64;

Query Match 31.4%; Score 750.5; DB 2; Length 486;
Best Local Similarity 35.0%; Pred. No. 8e-37;
Matches 171; Conservative 103; Mismatches 186; Indels 29; Gaps 9;

QY 3 KEVQNVNGEMKSSV--NQETISPIDDS-IGFVPMRREVDHAKGREALPAMAL 59
DB 7 QSYLNFINRMKSSSGETKEVLPADKDVGVQRTTIDVNEAIAKAKESWRKL 66
QY 60 TYVERAQLAKKADILIERDKERIATVLAKEISKAVNASTVEVTRADLIRYAEERIR-- 117
DB 67 SGNREGNLYLVANILLESRLDDVAETLTREMGKTLPEAKGFTARGAIIIRYAAGBLREV 126
QY 118 --LSTSADEGKMDSTGHKLAIVIRQPYGIYLAIPVYPVNLSGSKTAPALIGNVVM 175
DB 127 GDVISTDSGIMYTT-----RTPLGVGVITPMNPPVAPIPMKAPALIYNTVV 177
QY 176 FKPPGGSVGLVAKAPAPAGLPAGVFTITGRSGEIDGYIEHEVNFINTGSTPVG 235
DB 178 ITPATEAATAKAKVECFADAEIPDSVVMVVTGSGIVGNRIAEHDVAVGVTGSDQTG 237
QY 236 QRIKLAGNR--PIMELGKDGAGIVLADADLDNAKQIVAGAYDSGORTAIKRLVV 293
DB 238 KHIGGALAKGAKYQLEMGKKNPVIVADADLDLAKATISGSLSTGKCTATSVIYG 297
QY 294 EYVADELAEKISBNVAKLSVGPDPDNAT--VTPIYIDNSADPIESLVVDARQKAKEINLF 352
DB 298 RSVYDAFKEKLTQLQALITVGNGLBEDGNMNGSPVSEQLTVAHYIKKKEKATLLTGG 357
QY 353 KR-----DGRLLTPGFLFDHTLDMKLAEBEPFGILPIIRKDAEAVATANKSDGL 405
DB 358 HRLTDAEHKGNFVEFTVEDVNDMTIAQEBTFGVLLALIKVETBEAMKLANDTEYGL 417
QY 406 QSSVPTRDFOKAFDIANKLEVGTVINNTKGRGPNPFGLKGSQ--AGVQGIKRSIA 463
DB 418 SAAVFTENIRKMLTFIDIEVGLVVRNATAGVELAPFGKMKSSSHREGG-KAIIIF 476
QY 464 MTNYSIVL 472
DB 477 FTSIKTVFI 485

RESULT 43
Q97XS9 PRELIMINARY; PRT; 470 AA.
AC Q97XS9;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase, NADP dependent (GapN-1) (EC 1.2.1.9).
DE Name=gapN-1; OrderedlocusNames=SSO1629;
GN Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RA MEDLINE=21332296; PubMed=11427726; DOI=10.1073/pnas.141222098;
RA She Q., Singh R.K., Contaloniieri F., Zivanovic Y., Allard G.,
RA Awaraz M.U., Chan-Weiner C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Eranuo G., Fletcher C., Gordon P.M.K.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
RT "The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2";
RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006776; AAK41843.1; -.
DR PIR; D90323; D90323.
DR HSP; 057693; 1KY8.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde_dehydrog.
DR Pfam; PF00171; Aldehyd_1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 470 AA; 51529 MW; 54BE97B2D2FE78 CRC64;

Query Match 31.3%; Score 750; DB 2; Length 470;
Best Local Similarity 36.6%; Pred. No. 8.2e-37;
Matches 171; Conservative 94; Mismatches 186; Indels 16; Gaps 7;

QY 20 IEIISPIDDSLGGVPMRREVDHAKGREALPAMALTYVERAQLAKKADILIERDK 79
DB 7 IEVRSPNLTNYIGVYKRNHKEVGEIEAVKGPETISKMLPKRTAILRYSILEREQ 66
QY 80 EELATVLAKEISKAVNASTVEVTRADLIRYAEERIR--GIRLSTSADEGKMDA-----STG 133
DB 67 ERLATLAMEGKPIRDSRVVETLASRLFRHAAEEVGVLV---CKNVRVDAIYEPDNE 123
QY 134 HKLAIVIRQPYGIYLAIPVYPVNLSGSKTAPALIGNVVMFKPPTGSSVGLYAKAF 193
DB 124 NRIIVSTREPIGVTAIIPFNPINSFAHKAAPALAVNSVVKPSINTPLAATIMKIL 183
QY 194 AENGIPAGVFTITGRSGEIDGYIEHEVNFINTGSTPVGORTKLA---GMRPITLE 250
DB 194 VEAIGIPDSAVRVVYGSSEIGDEIITHPVLGLITLTGSLGKLASVAVSLGR-LIME 242
QY 251 LGGKAGAGIVLADADLDNAKQIVAGAYDSGORTAIRVAVVEVADELAEKISENVAK 310
DB 243 LGGSPITILIEBANIERASSIAVRAFPYACQNCAGKRIIVREIYDKFYKAFVDKARA 302
QY 311 LSVGDPDNAT--VTPIYIDNSADPIESLVVDARQKAYE--LNEFKRDDRLITPGLPHV 367
DB 303 LKVGDPLDETTDVGIVINKESVENINSVLADAKAKGKHYEILINKPESGSFFPLTMVNP 362
QY 368 TLDMKLAMEBPFGILPIIRYKDAEAVATANKSDPGLQSSVPTRDFOKAFDIANKLEV 427
DB 363 SLDMKLKSEVFGPIVIVSVKSDBEAIRINSTEYGLQSAIFINDVAPRLKLSRELKFG 422
QY 428 TVHINNTKGRGPNPFGLKGSAGVQGIKRSIAEAMTNVSIYVDM 474
DB 423 AVIINDSTRKWDLSLAFGPKKSSIGREGVAETMELENTLAIATL 469

RESULT 44
Q8CV96 PRELIMINARY; PRT; 497 AA.
AC Q8CV96;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Aldehyde dehydrogenase.
DE OrderedlocusNames=OB0861;
GN *Oceanobacillus theyensis*.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Oceanobacillus*.
OX NCBI_TaxID=182710;
RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=HTE831;
 RX MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkf526;
 RA Takami H., Takaki Y., Uchiyama I.;
 RT "Genome sequence of *Oceanobacillus theyensis* isolated from the Iheya
 RT Ridge and its unexpected adaptive capabilities to extreme
 RT environments";
 RL Nucleic Acids Res. 30:3927-3935(2002).
 DR EMBL: AP004596; BAC12817.1; --
 DR HSSP: P05091; 1004.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyd_dehydrog.
 DR Pfam: PF00171; Aldehyd_1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 497 AA; 54140 MW; 14F219B301BE54FB CRC64;

Query Match 31.0%; Score 742; DB 2; Length 497;
 Best Local Similarity 34.8%; Pred. No. 2, 6e-36;
 Matches 169; Conservative 98; Mismatches 196; Indels 22; Gaps 6;

QY 2 TBYQNYVNGEMSSVNOEILSPIDSSLGFPVPMATREVDHAKAGREALPAMAALTV 61
 DB 8 TDVLYKMFIDGKWDVNNNTTAVTNPAKERIVQVPLSDQSHVEAVALKAQKEMALVPA 67
 QY 62 YERAOYLAKRADIIERKKEIATVLAKESKAVNASVTEVTRADLIRYAABEGIRLSTS 121
 DB 68 PQRAEVLAVGMIMCKOKERLSRLTMENKVLBEARGEVQESIDMAFMAGGRLLF-- 125
 QY 122 ADEGKMDSTGKLAIVIRQVGYLAIPYVNVLSGSKAPALIGNVNMFEPPIQ 181
 DB 126 ---GQTPPAELKDFKMSQRPVGVGIIITPMNFPIAIVTKSPFAIVAGNVVMPATE 182
 QY 182 GSVGLVLAFAFAEAGLPAGVNTITGRGSEIGDYVEHEVNFITGSTPGVIGL 241
 DB 183 TPIMAYELKFIIEAGLPAGVNVVFGKSEVDSVQHKDIRKVISFTSNDTGRNIASD 242
 QY 242 AG--MRPIMLELGGKAGIVLADADLDNAKQIVAGAYDSGQRTAIRVAVEEVADE 299
 DB 243 CGKQKLKSVLEWGNKNAVIMDDADLDLAVEGLIWAFFGSGRCCTACSRVIVHKKVT 302
 QY 300 LAKISENVAKLSVGPDP--DNATVTPVIDNSADPIESLVDAQKAKELN----EFK 353
 DB 303 LBSRLAANDHLITGNGLEDGKIVGPIINEAGINKIKYIEVGKQAGATLLAGTEMETN 362
 QY 354 RGRLLTLPGLFDHVTLDMLAMEERGPILPIIRVDAEEVAVALAKSDPGLSSVFTPD 413
 DB 363 NPGYFTPLFTDAKPMRIADIEIFGPVSLIEVSSFEEALVNNQVTFGLSSSIFTXD 422
 QY 414 FOKAFDIANKLEVGTVHINKTGRGPDNPFPLGLKSG-----AGVQIRYSIEAMTNVK 468
 DB 423 VNRVFAQKQDLDIGIYVNAAGTTGAEIHLPFGSTKGTGNGHRDSGQ----ALDVTTEMK 478
 QY 469 STYLD 473
 DB 479 AIYVD 483

RESULT 45
 027059 PRELIMINARY; PRT; 455 AA.

ID 027059;
 AC 027059;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 GN NADP-dependent glyceraldehyde-3-phosphate dehydrogenase.
 OS OrderedLocustNames=MT978;
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 NCBI_TaxID=187420;

RP [1]
 RC STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Rother B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Carnso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pletrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of *Methanobacterium thermoautotrophicum*
 RT deltaH: functional analysis and comparative genomics";
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL: AE000871; AAB85474.1; --
 DR PIR: H69230; H69230.
 DR HSSP: O59931; 1EUH.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyd_dehydrog.
 DR Pfam: PF00171; Aldehyd_1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 455 AA; 48825 MW; A98D7BA54AA9C08B CRC64;

Query Match 31.0%; Score 741.5; DB 2; Length 455;
 Best Local Similarity 38.2%; Pred. No. 2, 5e-36;
 Matches 173; Conservative 83; Mismatches 186; Indels 11; Gaps 5;

QY 9 VNGEMSSVNOEILSPIDSSLGFPVPMATREVDHAKAGREALPAMAALTVYEAQVL 68
 DB 5 IDGEGVSGKIFTVNPFNGDEVDRVPLAGRSDEVALRAARADAMDLARKISEKL 64
 QY 69 HKAADIIEKKEIATVLAKESKAVNASVTEVTRADLIRYAABEGIRLSTSADEGCKA 128
 DB 65 YVADBLKLEDEPRLITLESQKPIRBSRDEKVSVEFARLSAEAGL---YGESIM 121
 QY 129 DASTGKH--LAVIRQPVGIVLAIPYVNVLSGSKAPALIGNVNMFKEPTQSVSG 186
 DB 122 DAGIGKGLTGTVAIPLGVAIATPPNYPLMIAHKVPALAGNTTIKPSLEAPLSA 181
 QY 187 LVLAFAFAEAGLPAGVNTITGRGSEIGDYVEHEVNFINTGSTPGVQRIQKLAGMR 246
 DB 182 LKLAITLSE--HPPAAVNAVATGRGSEVGVIIIDSPVLDKITTPGSEVGRYSIARSMTK 240
 QY 247 IMLELGGKAGIVLADADLDNAKQIVAGAYDSGQRTAIRVAVEEVADELAEKISE 306
 DB 241 ITLELGGNDPLVMDADIDSVEAAVGRSYISQVCLAVGRMIVHEVADAEFADKLVN 300
 QY 307 NVAKLSVGPDPDNAT--VTPVIDNSADPIESLVDAQKAKELNFEKDGRLTPGLFD 365
 DB 301 INGSLRAGDPMDVRIDVGPLINEDAIEVERVIGAAVEDGAILLCGSGRGNFVEPTVD 360
 QY 366 HTYLDMLKAMEEPGPILPIIRVDAEEVAVALAKSDPGLSSVFTPDQKAFDIANKLE 425
 DB 361 HVPQMEVIERETFGVSPILIRFNGADEAIRANGCYALQGVTEINRTLRMARREIE 420
 QY 426 VGVTHINKTGRGPDNPFPLGLKSGAGVQGR 458
 DB 421 AGTIVLVKQSTFRVDHMPF---GDSGAVENWR 449

RESULT 46
 08NM80 PRELIMINARY; PRT; 484 AA.

ID 08NM80;
 AC 08NM80;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 GN NAD-dependent aldehyde dehydrogenases (EC 1.2.1.-).
 OS OrderedLocustNames=C912668;
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 NCBI_TaxID=187420;

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OC NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;
 RA Nakagawa S.;
 RT "Complete genomic sequence of *Corynebacterium glutamicum* ATCC 13032.";
 RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP005282; BAC00062.1; -.
 DR HSSP: P51977; 1BXS.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyd_dehydrog.
 DR Pfam: PF00171; Aldehyd_1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 DR Complete proteome; Oxidoreductase.
 SQ SEQUENCE 484 AA; 51084 MW; A2E416EE6585775F CRC64;

Query Match 31.0%; Score 741.5; DB 2; Length 484;
 Best Local Similarity 36.9%; Pred. No. 2.7e-36;
 Matches 174; Conservative 93; Mismatches 184; Indels 21; Gaps 6;

QY 6 QNYVNGE-----KSYNQIILSPIDDSLGFPVPMTRBEVDHAMKAGREALPAAAL 59
 DB 12 KHLIGQWEGNSDRISTN---INPYDSVIAESKQSIADVDAYEAKKRAQAEWMAT 67
 QY 60 TYVERAQLYHKAADIIRDKELIATVLAKEISKAVNASVTEVVRADLIRYAEGRIS 119
 DB 68 PAAESATITVRAELIEHEHEBEIIVEMLIKESGTSKALEITLAGNTKESASPPGRV 126
 QY 120 TSADGCKMDASTGKHLAVIRROPVGIYLAIPYVNVLSGSKIALPALIGNVVFKXP 179
 DB 127 ---HGRISPSPTPGKENRVYVAVGVGVISPMNPELNLSTRSAVAPALAVGNVAVI 182
 QY 180 TQGSVSGVYL-AKAPAEAGLPAGVNTTIGRGEISGDIYVEHEVNFINTGSTPYGQRI 238
 DB 183 SPTPTGVGIPARIEEAGVPAGVISTVAGASGEIGDHFTVAVPEKLISFTSTPVGRRV 242
 QY 239 GKLA-----GMRPIMELGSKAGIYLAADLDLNAKQIYAGAYDVSQGRCTAIKVLVVE 294
 DB 243 GELALINGGPMKTVALELGNAPFVVALADIDPAAQAAVAPLHQGOICMSINRYIYDA 302
 QY 295 EVADLEIAEKISENVAKLSVGDPEPDNAT-VTPYIDNSADPIESLVVDARQKAKEINEFK 353
 DB 303 AVHDFLEKFEVAVAKNIPTGDSASGTLGVPIVINSQSLGELKEIKAKGATVQVEGR 362
 QY 354 RDGALLTPELPHVTLDMLAMEBPPGPILPPIIRVKDAEAVAIANKSDPGLQSSVFTRD 413
 DB 363 IEGRLVHPVFEVDVTSDMELIAREEIFGPIISYLKADDEHAAELANASDPGLSAAVWSD 422
 QY 414 FOXADIANKLEVGTVHINNTKGRGPDNFPPIGLKSGGAGVGIRYSIEAMT 465
 DB 423 IDRAAQFALQIDSGMVHINDLTVNDEPHVMFGSKNSGLGRNGWALIEEFT 474

RESULT 47

Q9HUR4 PRELIMINARY; PRT; 489 AA.
 AC Q9HUR4;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Probable aldehyde dehydrogenase.
 GN Ordered locus names=PA4899;
 OS *Pseudomonas aeruginosa*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardie K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Mong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004903; AAC08284.1; -.
 DR PIR: A83033; A83033.
 DR HSSP: Q63639; 1B19.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyd_dehydrog.
 DR Pfam: PF00171; Aldehyd_1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 DR Complete proteome.
 SQ SEQUENCE 489 AA; 52697 MW; DECBAA8E6E73793D8 CRC64;

Query Match 31.0%; Score 741.5; DB 2; Length 489;
 Best Local Similarity 37.0%; Pred. No. 2.8e-36;
 Matches 173; Conservative 87; Mismatches 185; Indels 23; Gaps 7;

QY 11 GEWK--SSVNOIEILSPIDDSLGFPVPMTRBEVDHAMKAGREALPAAALTYVERAQL 68
 DB 16 GEMHGRGRGRRLKXSNPFDGSLLEIEQADDDIDAAYAKAAEFQPAALGPSRAAVL 75
 QY 69 HKADIIRDYBEIATVLAKEISKAVNASVTEVVRADLIRYAEGRISTSA---DE 124
 DB 76 YKAVEVFPDRRHEIYDWTIRRESG-----TRLKAIEIWMGAARATLLESASPAPRVH 126
 QY 125 GKMDASTGKHLAVIRROPVGIYLAIPYVNVLSGSKIALPALIGNVVFKXPQGSV 184
 DB 127 GRIVSDVPKGSERYRRAICVGVISPMNPELHLTQRIAPALAGAAVVKKPSDTFV 186
 QY 185 -SGVLAFAFAEAGLPAGVNTTIGRGEISGDIYVEHEVNFINTGSTPYGQRIKLAG 243
 DB 187 CGGLLAIAIFEEAGLPAGLPFSVVGPSGEIDAVTEHHPGLVFTSTPVGRIKGLAS 246
 QY 244 ---MRPIMELGSKAGIYLAADLDLNAKQIYAGAYDVSQGRCTAIKVLVVEVADE 299
 DB 247 GGAHLKHALVEELGNSPVLVGLDADLEQAVNAVAFGKLEHQGOICMAINRIIVESDLYDA 306
 QY 300 LAEKISENVATLSVDP--PDNATYTPYIDNSADPIESLVVDARQKAKEINEFKRGR 357
 DB 307 PAARVEREVKGLRVGDPQPRADTA-VGPIVNAKQLEGLLEKIRLAKQESAKPLYESGVQ 365
 QY 358 LITPELPHVTLDMLAMEBPPGPILPPIIRVKDAEAVAIANKSDPGLQSSVFTRDPOKA 417
 DB 366 LIAHPVFEVATVMEIADLIFGPIVGLLRARDEHAAELANASDPGLSAAVFSRLERA 425
 QY 418 FDIANKLEVGTVHINNTKGRGPDNFPPIGLKSGGAGVGIRYSIEAMT 465
 DB 426 VFARQLRAQMTVHVDIPVNDENAPFGEKNSGLGRNGWALIEEFT 473

RESULT 48

Q6M2H6 PRELIMINARY; PRT; 496 AA.
 AC Q6M2H6;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DB BENZALDEHYDE DEHYDROGENASE (EC 1.2.1.28).
 GN Name=xy1C; Ordered locus names=CG2953;
 OS *Corynebacterium glutamicum* (Brevibacterium flavum).
 OC *Corynebacterium glutamicum*; Actinobacteridae; Actinomycetales;
 OC Bacteria; Actinobacteria; Corynebacteriaceae; Corynebacterium.
 NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 13032 / DSM 20300 / NCTB 10025;
RX PubMed=12994826; DOI=10.1016/S0168-1656(03)00154-8;
RA Kallinowski U., Bache B., Bartels D., Bischoff N., Bött M.,
RA Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Galgati L.,
RA Goemann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Pühler A.,
RA Rey D.A., Rieckert C., Rupp O., Salm H., Wendisch V.F., Wiegrebe I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins.";
RL J. Biotechnol. 104:5-25(2003).
DR EMBL: BX927156; CAP20692.1; -.
DR HSSP: P56533; 1A4S.
DR GO: GO:0018479; F:benzaldehyde dehydrogenase (NAD+) activity; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002066; Aldehyd_dehydrog.
DR Pfam: PF00171; Aldehyd_1.
DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 496 AA; 52386 MW; C4EB32B571BE3A84 CRC64;

Query Match 31.0%; Score 741.5; DB 2; Length 496;
Best Local Similarity 36.9%; Pred. No. 2.8e-36;
Matches 174; Conservative 93; Mismatches 184; Indels 21; Gaps 6;

QY 6 QNYVNGEM-----KSSVQNIELSPIDSSLGFPVPMTRREVDHMKAGREALPAMAA 59
DB 24 KHLIGQWEGNSDRISTN-----INPYDSVIAESQASADYDAVEAAKQAQAAAMAT 79
QY 60 TYVERAQYLAKADIIERDKESIAVLAKESIKAYNASVTEVVRADLIRYAAEGIRLS 119
DB 80 PAERSAIIYRAALEHEHEEIEVWLKESGSTRKAMLEITLACNITKESASPGRY- 138
QY 120 TSADEGKMDSGHLAVIRPOVGIYLAIPYNPVNLSSGSKIALPALLIGVNMFKPP 179
DB 139 ----HRIISPSTNGKENRYRAKGVGVISPMNPLNLSIRSAVALAVGAIVVICKA 194
QY 180 TQGSVSGLVL-AKAFPAAGLPAGVFNTITGRGSEIGDYVEHEBNVFINFTGSTPVGORI 238
DB 195 SDTPVGVGVIAPAFIEBAGVPACVISTVAGAGSEIGDHVTHAVPKLISITGSTPPGRVY 254
QY 239 GKLA-----GMRPIMLEGGKAGIVLADLDLNAKOIVAGAYVSGQRCTAIKRVLYVE 294
DB 255 GEIAINGGEMKTVALEGNAFPVVLADADIDAAQAAGAFIHOGQICMSINRYIVDA 314
QY 295 EVADELAEKISENVATLSGDDPPDNAT-VTPVTDNSADPIESLVVDARQKAKELNERK 353
DB 315 AVHDEFLERFVEAVKNIPFGDPSAEGTLVGVPVINDSQLKKEIKELAKKEGATVQVEGP 374
QY 354 RDRRLTLPGFVDVATLDMKLAMEEPFGPIIPIRVXAEEAVALANKSDFGLOSSVFTRD 413
DB 375 IEGRLVHPVFSVYTDMEIAREEIFGRLISVTKADDEAHAALANSDFGLSAAVWSKD 434
QY 414 FOKAFDIANKLEVGTVHNNKTRGPDNPFFLGKSGAGVQGIIRYSIEAMT 465
DB 435 IDRAAGFALQIDSGVHINDLTVNDPWHVFGSSKNSGIGRFNGDAIEEFT 486

RESULT 49
Q65NKO
ID Q65NKO PRELIMINARY; PRT; 488 AA.
AC Q65NKO;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE ycbd (Aldehyde dehydrogenase, Aldehyde dehydrogenase).
GN Name=ycbd; ORFName=BL01646; BL00285;
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxId=279010;

RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13;
RX PubMed=15383718;
RA Velth B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumer S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
RA Berka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species";
RT Genome Biol. 5:R77-R77(2004).
RL EMBL: AE017333; AAU39244.1; -.
DR EMBL: CP000002; AAU21892.1; -.
SQ SEQUENCE 488 AA; 52682 MW; 14F3C8F07D2D164F CRC64;

Query Match 30.9%; Score 739.5; DB 2; Length 488;
Best Local Similarity 35.3%; Pred. No. 3.6e-36;
Matches 171; Conservative 102; Mismatches 190; Indels 21; Gaps 9;

QY 2 TREYQNYVNGEM--KSSVQNIELSPIDSSL-GFPVPMTRREVDHMKAGREALPAMAA 58
DB 7 SKTYNFINGRWKVAESGMEOSLNPADTRDVLGVKXSIEDVDRAVEAAKQAQAAAMK 66
QY 59 LTVYERAOYLAKADIIERDKESIAVLAKESIKAYNASVTEVVRADLIRYAAEGIRL 118
DB 67 LGAERGGFLYKADIMEQRDEIAFCATREMGKTLPEAKGRTAIGTALIRYVAGGLRK 126
QY 119 STSADGKMDSGHLAVIRPOVGIYLAIPYNPVNLSSGSKIALPALLIGVNMFKPP 178
DB 127 T-----GDVIPSSTDSAPFYTRVPLGVGVISPMNPLNLSIRSAVALAVGAIVVICKA 181
QY 179 PTQGSVSGLVL-AKAFPAAGLPAGVFNTITGRGSEIGDYVEHEBNVFINFTGSTPVGORI 238
DB 182 ATEIAVTLKLVISCFEEAGIPSGVNAVTPSSSGQRIAEHPDVNGITFTSSNOTGKTI 241
QY 239 GKLAGNR--PIMLEGGKAGIVLADLDLNAKOIVAGAYVSGQRCTAIKRVLYVE 296
DB 242 GRTAFERGAQYQLEMGKRPVIVADADIDAVEAVISGAFSTGKCATSRVIVLNGV 301
QY 297 ADELAEKISENVATLSGDDPP-DNATVTPVTDNSADPIESLVVDARQKAKELNERK 349
DB 302 YDRPFEXKLQOTKEITIGDSLKEDVWGMPIANKQOLDNCLSIYANGKQAGADLIFGGERL 361
QY 350 NEFK-RDGLRLPGFVDVATLDMKLAMEEPFGPIIPIRVXAEEAVALANKSDFGLOSS 408
DB 362 ADGKTENGYYIRPAIFDNTVSGMTIAQEIIPVIALIADTLEELSTFANVVKRLSAS 421
QY 409 VFTRDFOKAFDIANKLEVGTVHNNKTRGPDNPFFLGKSGAGVQGIIRYSIEAMTN 466
DB 422 IFTQNRRLSTFTDEIENGLIRVNASAGVELQAFPGVQKSSSHSREG-BAAKEFFTA 480

RESULT 50
Q8U2S5
ID Q8U2S5 PRELIMINARY; PRT; 506 AA.
AC Q8U2S5;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase.


```

RC STRAIN=NCIMB 9866;
RX MEDLINE=95014081; PubMed=7929007;
RA Kim J., Fuller J.H., Cecchini G., McIntire W.S.;
RT "Cloning, sequencing, and expression of the structural genes for the
RT cytochrome and flavoprotein subunits of p-cresol methylhydroxylase
RT from two strains of Pseudomonas putida.";
RL J. Bacteriol. 176:6349-6361(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 9866;
RX MEDLINE=20029269; PubMed=10565539;
RA Cronin C.N., Kim J.-H., Fuller J., Zhang X.-P., McIntire W.S.;
RT "Organization and sequences of p-hydroxybenzaldehyde dehydrogenase and
RT other plasmid-encoded genes for early enzymes of the p-cresol
RT degradative pathway in Pseudomonas putida NCIMB 9866 and 9869.";
RL DNA Seq. 10:7-17(1999).
DR EMBL; U96338; AAA75634.2; -.
DR FMR; T46684; T46684.
DR HSSP; O63639; 1B19.
DR GO; GO:0016491; P:oxidoreductase activity, IEA.
DR GO; GO:0008152; P:metabolism, IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd, 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KM Plasmid.
SQ SEQUENCE 491 AA; 53156 MW; 70C45CE421266344 CRC64;

Query Match 30.8%; Score 736.5; DB 2; Length 491;
Best Local Similarity 35.8%; Pred. No. 5,5e-36;
Matches 171; Conservative 84; Mismatches 197; Indels 25; Gaps 7;

QY 5 YON---YNGEWK--SSVNOQIILSPIDSSLGFPVPMATREEDHAMKAGREALPAMAA 58
DB 8 YENMSIQLIAGEWRVKGAGDLDPDFTQEKLIQIPLAREDLDEAVSARQAQVAMAA 67
QY 59 LTYBERAOYLHKAADIIERDKEXIATVLAKEISKANASTEVNRRADLRVAEGIRL 118
DB 68 CGSEBAQVNLNVRIFDERRDIITWIIRESGS-----TRKAQIEWGABARATIQ 118
QY 119 ST----SADEGKMDSTGSHKLAVIRROPVGIATAPVNYPNVLSGSKIAPALIGANY 174
DB 119 ESMASLSRVAHGRILADVPKESRVRREPVGITISPMFPLHARSLAPLALGNAC 178
QY 175 MFKPRQGSVS--GLVIAKAPAEAGLPAGVFNITTGSGEIGDYIVHEBEVNFINTGSP 233
DB 179 VIKPASDTPYTGGLLAAHIFEEAGLPKGVLSVVVGSGSEIGDAFVHEHVPGFISFTGS 238
QY 234 VQGRIGKLAG---MPPIMELGSKAGIYLAADDDNNAKQIVAGAYDYSQGRCTAIR 289
DB 239 VGNRIIRIAGGHEHLKVALIELGNSPPVVLADADIDQAVNAVVGKFLHOGQICMAI 298
QY 290 VLVVEEVADELAKEISENVAKLSVGPDPMAVT--PYIDNSADFIESTLVVDAROKA 348
DB 299 IIVEDSVYDFVNRVYKRVKSLPYGDPKREIVVGVYINAKQLAGIQDKIATKSGAR 358
QY 349 LMEFKDGRLLTPEGLPDHYTLDMKLMWBEPPGPIILIRYKADEBAVALANKSPFGL 408
DB 359 MYEGAGQGVNLPHPVADVTADMEIAREEIFGRLVGIQRRARDAAHLELANSEYGL 418
QY 409 VFRDQKAFDIAKLEVEGTVAHNNKTGKQPNFPPLGLKSGAGVQGIIRYSIEANT 465
DB 419 VFTSLERGVKFAIRGAGTHTINDIPVDEPNAPFGEGKSGKIGRFGMDWAIIEPT 475

RESULT 53
O81OX6 PRELIMINARY; PRT; 474 AA.
AC O81OX6; O6H242; Q6KTA1;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Aldehyde dehydrogenase family protein.

```

```

GN OrderedLocusNames=BA2289, BAS2135, GBA2289;
OC Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Bailly L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Oksrad O.A., Helgason E., Ristone J., Wu M.,
RA Kolony J.F., Beaman M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Kouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plant R.D., Wolf A.M., Watkins K.L., Nierman J.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaitte J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related Bacteria.";
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Kasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Stemne;
RA Bretzin T.S., Bruce D., Chailacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RL Submitted (JAN-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; AE017031; AAP26160.1; -.
DR EMBL; AE017334; AAT31413.1; -.
DR EMBL; AE017225; AAT54447.1; -.
DR HSSP; Q28399; 109J.
DR TIGR; BA2289; -.
DR TIGR; GBA2289; -.
DR GO; GO:0016491; P:oxidoreductase activity, IEA.
DR GO; GO:0008152; P:metabolism, IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd, 1.
KM Complete proteome.
SQ SEQUENCE 474 AA; 51998 MW; 70EC5347DC9D22D CRC64;

Query Match 30.8%; Score 736; DB 2; Length 474;
Best Local Similarity 35.7%; Pred. No. 5,7e-36;
Matches 168; Conservative 99; Mismatches 195; Indels 8; Gaps 5;

QY 8 YNGEKKSVNOQIILSPIDSSLGFPVPMATREEDHAMKAGREALPAMAAITVYRAQ 67
DB 6 YINGMKSVNYTKPYAPVASEETLARIAGTEDEYKEAANAAMKKNNTLSAYDRAT 65
QY 128 YNGEKKSVNOQIILSPIDSSLGFPVPMATREEDHAMKAGREALPAMAAITVYRAQ 185
DB 123 LDPAPGADRIAYTRKPIGVIGALTTPFPLNVAAHKVGPALAGNTVLLPADQTP 182
QY 68 LKRAADIIERDKEXIATVLAKEISKANASTEVNRRADLRVAEGIRLSTADEGK 127
DB 66 LKRVAKMDERREBEIITAKAAKPIPAARCEVDITVOTYFAAEAKRI---YGETIP 122
QY 128 MDASTG--HKLAVIRROPVGIATAPVNYPNVLSGSKIAPALIGNVVMPKPTQGS 185
DB 123 LDPAPGADRIAYTRKPIGVIGALTTPFPLNVAAHKVGPALAGNTVLLPADQTP 182
QY 186 GLVIAKAPAEAGLPAGVFNITTGSGEIGDYIVHEBEVNFINTGSPVQGRIGKLAG 245
DB 183 SYALIELFEAGLPGKALNITISGPGSTVGEALVTVDVASITFTSPKGIQIKAKAG 242
QY 246 PIMLEIGSDADIVLADADL--DNAKQIVAGAYDYSQGRCTAIRKLVVEEVADELA 304
DB 243 RVTLEIGSNAAVITIDEDVELTDLIERVAMGAFVANNQVCISQVRFVHEERDDPL 302

```

```

Oy 305 SENNAKLSVGP-PNNATVTPVIDNSADFISLVDAROQAKEL-NEFKDRGLLTG 362
Db 303 RKAMETVVGGPILEETDVSALSKVDRINDVQEAVEKANVLYGKKDERVEFT 362
Oy 363 LFDHTLTMKLAEEPPGILPIIRVKAEAVALANKSDFGLQSSVTRDRQKAFDIAN 422
Db 363 VLTNVPEHVSQQCVFPGFLMTVNTPKEDEAIEOVNNSRYGLQAVFNNLFKMRAID 422
Oy 423 KLEVTVAHNNKGRGPNFPELGKSGAGVQGIRSIEAMTNVKSIVL 472
Db 423 ELEVGVMVINDIPTFRVDHMPYGVKESGTGEIKYALEMTEMKLVCI 472

RESULT 54
Oy 06HJ76 PRELIMINARY, PRT, 474 AA.
Oy 06HJ76
AC 06HJ76;
DT 05-JUN-2004 (TREMblrel. 27, Created)
DT 05-JUN-2004 (TREMblrel. 27, Last sequence update)
DT 05-JUN-2004 (TREMblrel. 27, Last annotation update)
DE SuccinC-semaldehyde dehydrogenase NMD(P)+ (EC 1.2.1.16).
GN Name=gabd; OrderedLocustNames=B9727.2073;
OS Bacillus thuringiensis (subsp. konkukian).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=180855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-27;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K., Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R., Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus thuringiensis 97-27.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AE017355; AAT59797.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0009013; F:succinate-semialdehyde dehydrogenase [NAD(P) . .; IEA.
DR GO; GO:0008155; P:metabolism; IEA.
DR InterPro; IPR02086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldedh; 1.
DR Complete proteome; Oxidoreductase.
SQ SEQUENCE 474 AA; 51973 MW; 9F8C4A3C76AC678F CRC64;

Query Match 30.8%; Score 736; DB 2; Length 474;
Best Local Similarity 35.7%; Pred No. 5.7e-36;
Matches 168; Conservative 99; Mismatches 195; Indels 8; Gaps 5

Oy 8 YNGEWMKSSVNOQIEILSPIDSSLGFVPAMTREEDVDMAKAGREALPAMAALTVERAYOY 67
Db 6 YINGDMKSVNTYKPLYAPYSEETLABIAQGTEDVEAVAAKKAMKEMNTLISADRATI 65
Oy 68 LHKADDIIERDKEELATYLAKISKAYNASVTEVRTADLIRYAEEGRISTSADEGSK 127
Db 66 LEKVAQKNDERREEEFAEIIAKEAPRIRARGESVRTVTYFAAEAKRI--YGETLP 122
Oy 128 MDASTG--HKLAVIROPYGIVALTAYVPYPNLSGSKAPLIGSNVVMFKPPTGOSVS 185
Db 123 LDAAGALGRIAYTRKPIGVGALTPFRFPLVAHKGPMAAGNTVLPADTPIS 182
Oy 186 GLVIAKAPAEADLPAGVFNTIGRGSEIDYIVEEHVNFINFTGSTPVQRIGKLAGR 245
Db 183 SYALIELPEEAGLPGALNIISGPSTVEALVTNDDVASIFTGTSPKGIGIKAKAGIK 242
Oy 246 PIMELSGKDAGIVALADAD-DNAKQIYAGVADYSSGQCTAIKRYLVVEVADELAEKI 304
Db 243 RVTELEGNSAAVTIDEVELTELERVMGAFVNNGCISIVORFVAEBRMDPLSTL 302
Oy 305 SENVAKLSVGP-PNNATVTPVIDNSADFISLVDARQAKEL-NEFKDRGLLTG 362
Db 303 RKAMETVVGGPILEETDVSALSKVDRINDVQEAVEKANVLYGKKDERVEFT 362
Oy 363 LFDHTLTMKLAEEPPGILPIIRVKAEAVALANKSDFGLQSSVTRDRQKAFDIAN 422

```

DB	VLTVNPEHVS	QOCQEV	GPLMTVNTVFKEFDALIECVNNSRGLQAGVETNNLFC	KRA	RAID	422
QY	423	KLEVGVTHNNKTRG	GPDPNPFPLGKSGAGVQIGIRYSIEAMTVKSYIVL	472		
DB	423	ELFVGVWINDI	PTFRVDHMPYGGVKGSGTREGIKYVIEEMTEKLVCI	472		
RESULT	55					
Q738S1		PRELIMINARY;	PRT;	474	AA.	
AC	Q738S1					
DT	05-JUN-2004	(TREMBlrel. 27, Created)				
DT	05-JUN-2004	(TREMBlrel. 27, Last sequence update)				
DT	05-JUN-2004	(TREMBlrel. 27, Last annotation update)				
DE	Aldehyde dehydrogenase family protein.					
GN	OrderedLocusNames=BCE2323;					
OS	Bacillus cereus (strain ATCC 10987).					
CC	Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.					
OK	NCBI_TaxID=222523;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	PubMed=14960714; DOI=10.1093/nar/gkh258;					
RA	Raako D.A., Ravel J., Oekstad O.A., Helgason B., Cer R.Z., Jiang L.,					
RA	Shores K.A., Fouts D.B., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,					
RA	Neison W.C., Kolstoe A.-B., Frazer C.M., Read T.D.;					
RT	"The genome sequence of <i>Bacillus cereus</i> ATCC 10987 reveals metabolic					
RT	adaptations and a large plasmid related to <i>Bacillus anthracis</i> pXO1."					
RL	Nucleic Acids Res. 32:977-988(2004).					
DR	EMBL; AE017271; AA541241.1; -					
DR	HSSP; P20000; IAG8.					
DR	TIGR; BC82323; -					
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.					
DR	GO; GO:0008152; P:metabolism; IEA.					
DR	InterPro; IPR002086; Alddehyd_dehydrol.					
DR	Pfam; PF00171; Aldedh; 1.					
KM	Complete proteome.					
SO	SEQUENCE 474 AA; 52079 MW; DCC692BB045D00AF CRC64;					
Query Match	30.7%;	Score 735;	DB 2;	Length 474;		
Best Local Similarity	35.7%;	Pred. No. 6.5e-36;				
Matches	168;	Conservative 98;	Mismatches 196;	Indels 8;	Gaps 5	
QY	8	YNGEWSVNVQIEILSPIDSSLGFPVPMTRREVDHAMKAGREALPMAALTYTERAY	67			
DB	6	YINDWKSVMNTYKFLYAPYSEETLAEIQGTEEDVKEAVIAKNAKMEKNRLSAVDRAIT	65			
QY	68	LHKADIIERKEBIATYLAKEISKANYASTVEVTRTDLRVYAEBSIRLSTSDDEGK	127			
DB	66	LEKVAQKDERREFEPAETIAEAKPKIPRAAGEVDRTQYTKFAEEAKRI--YGETLP	122			
QY	128	MDASTG--HKLAIVRRPOVGIVLIAIPYNYEVNLSGSIAPALIGANVMFPKPTQGSVS	185			
DB	123	LDAAFGADGRVAYITRKPIGIVGALTTPNFPLNLVAHKVGPALIAAGNTVVLKPADQTPIS	182			
QY	186	GLVIAKFAEAGDPAGVFNTITGSGSEIGDIVHEHEVNFINTGSTPEVGRIKGLAGNR	245			
DB	183	SYALVELPEEAGLPKGLKALNTISGPGSTVGEAIVTNDVASTITFTGSPKVLGIKAKAGIX	242			
QY	246	PIMLELGGKDAIVLADADL-DNAKQIVAGAYDYSGGRCALIKRVLYVEEVADELAEKI	304			
DB	243	RVTLELGSNAVAIIIDEDVELTDELIERVKGMAFVNNCGVCISVGRVFAHEEEMDEFLSKL	302			
QY	305	SENVAKLSVGRP-FDMATVTFVVIDNSADTFESLVVDAROGAKEL-NFKRDRGLTPFG	362			
DB	303	QKAMETVVGGPLLEETDVSALISKQDVERIDNWQEAIVKGAASYLVGAGKSDERI FKPT	362			
QY	363	LPFDVTTDMKLAMESEFPGLPIIRIVKDAEEAVAIANKSDGLQSSVETRPDOKAFDIAN	422			
DB	363	VLTVNPEHVSQOCQEVGPLMTVNTVFKEFDALIECVNNSRGLQAGVETNNLFC	422			
QY	423	KLEVGVTHNNKTRG	GPDPNPFPLGKSGAGVQIGIRYSIEAMTVKSYIVL	472		
DB	423	ELFVGVWINDI	PTFRVDHMPYGGVKGSGTREGIKYVIEEMTEKLVCI	472		

RESULT	56
ID	09KC36
ID	09KC36 PRELIMINARY; PRT; 482 AA.
AC	09KC36
DT	01-OCT-2000 (TREMBLrel. 15, Created)
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase.
GN	Ordered locusNames=BH1738;
OS	Bacillus halodurans.
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC	NCBI_TaxId=86665;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C-125;
RX	MEDLINE=20515282; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA	Takami H., Nakasome K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA	Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kihara S.,
RA	Horikoshi K.;
RT	"Complete genome sequence of the alkaliphilic bacterium Bacillus
RT	halodurans and genomic sequence comparison with Bacillus subtilis.";
RL	Nucleic Acids Res. 28:4317-4331(2000).
DR	EMBL; AP001513; BAB05457.1; ..
DR	PIR; B83867; B83867.
DR	HSSP; Q5931; 1EUN.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0008152; P:metabolism; IEA.
DR	InterPro; IPR02086; Aldehyde_dehydrog.
DR	Pfam; PF00171; Aldedh; 1.
KW	Complete proteome.
SO	SEQUENCE 482 AA; 52424 MW; 29DFBCABE1F8D4D CRC64;
Query Match	30.7%; Score 735; DB 2; Length 482;
Best Local Similarity	35.0%; Pred. No. 6.7e-36;
Matches 164; Conservative 90; Mismatches 209; Indels 6; Gaps 3	
QY	9 VNGEKSSVNOQIEILSPIDSSLSGVPAMTREEDVHAKGREALPMAALTYVERAOYL 68
DB	12 ISDQHQOQSIEVIDPGTNHVIATVPATEADAKAIEAAKEGAEIRAROMPHERAAIL 71
QY	69 HKAADIIRDKEEITVLAKESIKAYNASVTEVNTALDIRAAEEGRILSTSDGGM 128
DB	72 KRTAEIVTRDELFPAQTTAKESGKTIIRARSEVRRCIETLTLSGEARLHGETTIPFSOM 131
QY	129 DASTGH--KLAVIRROPVGIYLAIPYVYVNLSSGSKTAPALIGNVVMFKPPTGOSVSG 186
DB	132 ---PGEHRVGVGFYFPIGITVIAIPFPDPLNLVHKIGPALAAGNSLIRKSSPTPLSA 188
QY	187 LVYAKAFPEAGIPAGVFNNTTGRGSEIGDYVEHEEVNFINTGSTPVGQRIQKLAGMRP 246
DB	189 LRUVVAALEAGIPKKITVQITVGHGSVISPRTLKHDVRLISPTGGEYEGEKIARSGAYKK 248
QY	247 IMLEEGGDAGIVLADADLDNNAKQIVGAYDVSQRCATIRVIVVEBVDELAEKISB 306
DB	249 LAMELGNSPTITVLDDAEIMEAVASCVSGAFQAAGQNCIGVQRTIVBOSVFNSFIQSYVA 308
QY	307 NVAKLSVGDPEPDNAF-VTPVIDNSADEIESIIVADROKAGKALNEFKRDRGLTPGLFD 365
DB	309 QTKQRLDKQSEEDTIGPMISEKAKRIERWVEAKEGARVLIGGRITGAFFEPVILT 368
QY	366 HTYLDMLKAMEEPPGPIILPIIRYMDAEAVAIANKSDGLOSVSVTRPQKAFDLANTLE 425
DB	369 NVGPESRILAKEBAFAPIVYIRGVSHSLTEIAANADVDTGLQGLGTLNNLTAFAISAIETLE 428
QY	426 VGTVAHNKTKGRGPDNFPFLGLKSGAGVQGIYRSIEAMTVKYSIVLDM 474
DB	429 VGGIVWVNDSDVIRIDAMPFGGIIKTKSGLREGVRVAIEEMTEQKVAAVFL 477

ID	Q8TWC7	PRELIMINARY;	PRT;	471 AA.
AC	Q8TWC7			
DT	01-JUN-2002 (TREMBlrel. 21, Created)			
DT	01-JUN-2002 (TREMBlrel. 21, Last sequence update)			
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)			
DE	NAD-dependent aldehyde dehydrogenase.			
GN	OrderedLocustNames=MK1108;			
OS	Methanopyrus kandleri.			
OC	Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;			
OC	Methanopyrus.			
OK	NCBI_TaxID=2320;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=AV19 / DSM 6324 / JCM 9639;			
RX	MEDLINE=21927647; PubMed=11930014; DOI=10.1073/pnas.032671499;			
RA	Slesarev A.I., Mezhevaya K.V., Makarova G.S., Polushin N.N.,			
RA	Shcherbinina O.V., Shaknova V.V., Belova G.I., Aravind L.,			
RA	Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,			
RA	Mal'kh A.G., Koonin E.V., Kozlovskiy S.A.;			
RT	"The complete genome of hyperthermophile Methanopyrus kandleri AV19			
RT	and morphology of archaeal methanogens."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).			
DR	EMBL; AE010400; AA002321.1; -.			
DR	HSSP; A51977; IBXS.			
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.			
DR	GO; GO:0008152; P:metabolism; IEA.			
DR	Pfam; PF00171; Aldedh; 1.			
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.			
KW	Complete proteome.			
SQ	SEQUENCE 471 AA; 51147 MW; E369538CE2D7B3D CRC64;			
Query Match 30.7%; Score 734; DB 2; Length 471;				
Best Local Similarity 36.7%; Pred. No. 7.4e-36;				
Matches 173; Conservative 90; Mismatches 201; Indels 8; Gaps 5				
QY	4 EYQNVVNGEMKSVNQVQRIITLSPIDSSIGFVPMATREVDHMKAGRELPMALTYVE	63		
DB	2 EYGLIGEGMELEGREIIVENPNPDGSEVGRVLA-DEVDEALLKQAREGQRRRERTPYE	60		
QY	64 RAQYLHKRAADIERDKEEIAVTLAKEISKAVAVSVEVTRADLIRYAAEEGIRLSTSD	123		
DB	61 VRELAELAHAHLKHGHEDELAELIALEGKPIRDAYEYRTREVLRSAAEABRLYGETL	120		
QY	124 EGGKMDASTGH-KLAVIRQPVGIVALIAPYNEVNTSGSKIAPALLIGNVVMFKPQTQ	181		
DB	121 PG--DAQRGRTAEILITVREPVGVLSTIPPYNFPLLLPTHKLGPALARCSVHKPATV	177		
QY	182 GSVGSLVLAQFAERGLPAGVFNNTTGGSELEGDYIVHEEYNFNFTGSGTGVQRIQL	241		
DB	178 TPISLSRLAEILLDAGVPLALQVVVGGAEIGEBIA-RAPDPAISFTSGSRVGEHLREI	236		
QY	242 AGMPIMLELGKAGIYVLADADLNAAKOIYAGVYDVSQGRCTAIKRVLYVEVDELA	301		
DB	237 SPIRITITELGANDFVYDETRADVEAAAEAAVGRGCTHAGQVCIAVERAIVVEDVYEBEL	296		
QY	302 EKISENVAKLSVGPEDFNAT-VTPYVIDNSADFISSLVVDAROKAKAEINEFKRDGRLLT	360		
DB	297 EAAEVAEASLKVGPDLDEDTVGPIIDGVAEKVARRHEDAVERGAELITGSEBGRLLP	356		
QY	361 PGLFDHVTLLDKMLAEERPFGPIILPIIRYKDAEAVAIANKSDFGLOSSVFTRDFOKAFDI	420		
DB	357 PTVLADVPEDALVAREETFGEPVLPIYIRAKDEEBAIRINSTGYGLAAVFTERLDRAVXA	416		
QY	421 ANKLEVGIVHNNKTRGPRDNPPLGLTGSAGVGCIVYSIEAMNNVYSIVL	472		
DB	417 TRKLDAGVIVNESTIYRADYMPFGGVAASGVGRGVPQAAVEKFTBEKTVI	468		

DT 25-OCT-2004 (TREMblrel. 28, last sequence update)
 DT 25-OCT-2004 (TREMblrel. 28, last annotation update)
 DE Succinate-semialdehyde dehydrogenase NAD(P) + (EC 1.2.1.16).
 GN Name=gabd; ORNames=BTZK2069;
 OS Bacillus cereus ZK.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxId=288681;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ZK:
 RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Rice H.;
 RT "Complete genome sequence of *Bacillus cereus* ZK."
 RL Submitted (Jul-2004) to the EMBL/Genbank/DBJ databases.
 DR EMBL: CP000001; AU18186.1; -.
 KM Oxidoreductase;
 SQ SEQUENCE 474 AA; 51985 MW; 246863CBA616C48 CRC64;

Query Match 30.6%; Score 732; DB 2; Length 474;
 Best Local Similarity 35.7%; Pred. No. 9,9e-36;
 Matches 168; Conservative 98; Mismatches 196; Indels 8; Gaps 5;

QY 8 YVNGEWSVNOIEILSPIDSSLGFPVPMTRREVDHAKGREALPAMALTYERAC 67
 Db 6 YVNGEWSVNOIEILSPIDSSLGFPVPMTRREVDHAKGREALPAMALTYERAC 65
 QY 68 LHKADITRDKEEITATVAKESIKAVNVAETVETALINVAEGIRLSTSDGEGK 127
 Db 66 LHKVAKMDERREEFETIAKAPKIRARGSEVDVTOTYFPAEAKRI--YGETLP 122
 QY 128 MDASTG--HKLAVIRQPVGIYLAIPYNYPVNLSSKIPALLIGNVMPKPTQGSVS 185
 Db 123 IDAAGADGRIVYTRKPIGVIGALTPNPPLNVAKHKGPAAGNTVVLPAQOTPS 182
 QY 186 GVLAKAFPAEAGLPAGVNTTTRGSEIDYVEHEVNFINTGSTPVGORIGKLAGR 245
 Db 183 SYALIELPEAEGLPKGALNIGSPGSTGEALVTNDVASITFTGSPKIGIKKAKAGK 242
 QY 246 PIMLELGGADAGIVLADNL-DNAKQIVAGAVDSGQCTAIKRVLYVEVADAELEKI 304
 Db 243 RYTLIEGSNAAVIIDEVELTELERVWGAFVNNQVCISIVQRFVHEERNEFLSL 302
 QY 305 SENVAKLAVGDP-FDNAATVPYIDNSADPIESLVVDARQKAKEL-NFKRDGRLTPG 362
 Db 303 RRAMETVVVGDPDLEETDVSALISKDVRIDNWQEAKEGANLYGKAKDERVFEPT 362
 QY 363 LFDHVTLDKLAEEPPGPIILPIRYKAEAEVAVALANKSPGLQSSVFTRDFOKAFDIAN 422
 Db 363 VLTNPVEHVSQCQEVFGPLMTVNFKEFNEALIEQVNNSRYGQAGVFTNNLFFKMRALD 422
 QY 423 KLENGTVHNNKTRGQDPNFPPLGLKGSAGVOGIRYSTEANTVYSITL 472
 Db 423 ELEVGQVMINDIPTRFVDHPYGVKESGTGRBGIRYALEEMTEMLVCI 472

RESULT 59
 GABD_ECOLI
 ID GABD_ECOLI STANDARD; PRT; 482 AA.
 AC P25526; P78207; P78208; P78209;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 25-JAN-2005 (Rel. 46, Last annotation update)
 DE Succinate-semialdehyde dehydrogenase [NADP+] (EC 1.2.1.16) (SSDH).
 GN Name=gabd; OrderedLocusName=b2661;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxId=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / JMI03;
 RX MEDLINE=94127927; PubMed=8297211;

RA Niegemann E., Schulz A., Bartech K.;
 RT "Molecular organization of the *Escherichia coli* gabd cluster:
 RT expression of the GABA permease gene.";
 RL Arch. Microbiol. 160:454-460 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97349980; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474 (1997).

RL Science 277:1453-1474 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97349980; PubMed=9205837;
 RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Maki T., Mitsuhashi N.,
 RA Mizubuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
 RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubraman S.,
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
 RA Yamagata S., Horiuchi T.;
 RT "Construction of a contiguous 874-kb sequence of the *Escherichia coli*-
 RT K12 genome corresponding to 50.0-68.8 min on the linkage map and
 RT analysis of its sequence features.";
 RL DNA Res. 4:91-113 (1997).
 CC -1- CATALYTIC ACTIVITY: Succinate semialdehyde + NAD(P) (+) + H(2)O =
 CC succinate + NAD(P)H.
 CC -1- PATHWAY: 4-aminobutyrate (GABA) degradation.
 CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
 CC in positions 51 and 190.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M88334; AAC36831.1; -.
 DR EMBL: U00096; AAC3708.1; -.
 DR EMBL: D90890; BAA16522.1; ALT_FRAME.
 DR EMBL: D90890; BAA16523.1; ALT_FRAME.
 DR EMBL: D90890; BAA16524.1; ALT_FRAME.
 DR PIR: F65045; F65045.
 DR HSSP: Q28399; 1097.
 DR ECHOBASE; EB1305; -.
 DR Ecogene; EG11329; gabd.
 DR InterPro; IPR02086; Aldehyde_dehydr.
 DR Pfam; PF00171; Aldehyd; 1.
 DR TIGRFAMs; TIGR01804; BADH; 1.
 DR TIGRFAMs; TIGR01722; MMSDH; 1.
 DR TIGRFAMs; TIGR01780; SSADH; 1.
 DR PROSITE; PS00707; ALDEHYDE_DEHYDR_CYS; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
 KW Complete proteome; NADP; Oxidoreductase.
 FT NP BIND 233 238 NAD(P) (ADP part) (By similarity).
 FT ACT_SITE 255 258 By similarity.
 FT ACT_SITE 289 289 By similarity.
 SQ SEQUENCE 482 AA; 51720 MW; 091538F8741DB0CF CRC64;

Query Match 30.3%; Score 726; DB 1; Length 482;
 Best Local Similarity 34.3%; Pred. No. 2,3e-35;
 Matches 164; Conservative 92; Mismatches 206; Indels 16; Gaps 7;

QY 6 QNYVNGEWSVNOIEILSPIDSSLGFPVPMTRREVDHAKGREALPAMALTYE 63
 Db 12 QALINGEWLDANNGBAIDVTNPANGDKLGSVPMKAGDTRAIIDANRALPAMRALTAKE 71

QY 64 RAQVHLKADIIERDKERIATVLAKEISKAYNAVTEVTRADLIRYAAREGIRLSTSD 123
DB 72 RATTILNENLWMEHODDLARLMTLLEOQPLAEAKEISYAASFIEWAEERIGYDTI 131
QY 124 ECGKMDASTGKLAIVRRPQVGIYLAIPYVNVNLSSGIAPALIGNVVMPFPPTQGS 183
DB 132 PGHQAD-----KRLIYIKOPIGITTAITTPMNFPAAMITRKAGPALAACCTMVLKPSQTP 186
QY 184 VSGDLVLAKAPAEAGLPAGVFNITITGRGSEIGDYIVHEEVNFINFTGSTPVGORIGKLAG 243
DB 187 FALALAEIARAGVAGVFNVTGSAGAVGNELTGNPLVYKLSFGSTGTEIGRLMEQCA 246
QY 244 --MRPIMLELGGKAGIYVADADLDMAAQIYVAGVYSGQRCTAIRVLVVEVADLA 301
DB 247 KDIKRYKSLBGGAPPIVFDDADLDKAVGALASKFRMAGQTCVCANRLYVODGVYDRPA 306
QY 302 EKISENVAKLSVGPDPDN-ATVTPVIDNSADPIESLVDAROKAKEL---NEFRDR 357
DB 307 EKLQCAVSLKIHIDGDLNCGVTTIOPILDERKAVAKVEHIALDEKARVCGGKAHERGN 366
QY 358 LLTPGLFDHYVTLDMKLAWEPPGPIILPIIRVKADEAVAIANKSDPGLQSSVFTDFOKA 417
DB 367 FQOPTLVDPVPAKAKVSKETEGPLAPLFRFQDEADVIAQANDTERGLAAVFAARDLSRV 426
QY 418 FDIANKLEVGTVHINNTGRCQDPNPFPLGLKSGAGVQGIKRYSIEMTNVKSIVL 474
DB 427 FRVGALEXYGIVGIN--TGIIISNEVAPFGGIKASGIGREGSKYGLDYLEIKYMCIGL 482

RESULT 60

ID 072100 PRELIMINARY; PRT; 488 AA.
AC 072100;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Succinate-semialdehyde dehydrogenase (EC 1.2.1.16).
GN Name:sgad; OrderedlocusNames=LMO12365.0935;
OS *Listeria monocytogenes* (serotype 4b / strain F2365).
OC Bacteria; Firmicutes; Bacilliales; *Listeriaceae*; *Listeria*.
OX NCBI_TaxID=265669;
RN [1]
RP SEQUENCE FROM N.A.
RA PubMed=15115801; DOI=10.1093/nar/gh562;
RA Nelson K.E., Foute D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
RA Kolonay J.F., Rasko D.A., Anguoli S.V., Gill S.R., Paulsen I.T.,
RA Peterson J.D., White O., Nelson W.C., Nieman W.C., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA Haft D.H., Selengut J., Van Aken S.E., Khouri H.M., Fedorova N.,
RA Forberger H.A., Tran B., Kacharion S., Wondolilling L.D., Unlith G.A.,
RA Bayles D.O., Luchansky J.B., Fraser C.M.;
RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
borne pathogen *Listeria monocytogenes* reveal new insights into the
core genome components of this species.";
RL Nucleic Acids Res. 32:2386-2395(2004).
DR EMBL; AE017325; AAT03714.1; -.
DR HSSP; P20000; 1A68.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0009013; F:succinate-semialdehyde dehydrogenase [NAD(P) .]; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde_dehydrog.
DR Pfam; PF00171; Aldehyd. 1.
DR Complete proteome.
SQ SEQUENCE 488 AA; 53194 MW; ACE7841259C0D152 CRC64;

Query Match 30.3%; Score 726; DB 2; Length 488;
Best Local Similarity 34.2%; Pred. No. 2.3e-35;
Matches 163; Conservative 95; Mismatches 200; Indels 18; Gaps 6;
QY 8 YVNGEKSSVNO--LEILSPIDSSLGFPAMTRREVDHAMKAGREALPAMALTYVERA 65
DB 16 FINGKMTDNDNKETQDIYVAPNGEVYAKIAQAGPNETKKAIRAKAFAFDMAKMEADLV 75

QY 66 QYHLRAADIIERDKERIATVLAKEISKAYNAVTEVTRADLIRYAAREGIRLSTSD 125
DB 76 KLIHLKADIMEKADATLAKIMTLEOQPLKESKEGYLTGAENFRFAAEARL-----YG 130
QY 126 GMDASTGKLAIVRRPQVGIYLAIPYVNVNLSSGIAPALIGNVVMPFPPTQGS 185
DB 131 ETIPAPNNHAF-IVKQPIGVAAITTPMNFPGAMTRKLAIPALATGNTIVLPGSGDTPLS 189
QY 186 GLVLAAPAEAGLPAGVFNITITGRGSEIGDYIVHEEVNFINFTGSTPVGORIGKLAG-- 243
DB 190 ALAIEIFEIEEALPGVANNITWSSKEIGETLITDSODVAKLFTGSTKXGQLPFGQSAFT 249
QY 244 MRPIMLELGGKAGIYVADADLDMAAQIYVAGVYSGQRCTAIRVLVVEVADLA 303
DB 250 LKISLIELGHPPIVFDDANIDAVNDIAAKFRNNGCVCSPNKIFAKIKKEKFTYA 309
QY 304 ISENVAKLSVGPDPNATVTPVIDNSADPIESLVDAROKAKELNEFKR-----DG 356
DB 310 LVAKVBQLKVGGLDVNVGPIREDADIDKIDKOLKNATDKAKAVLTGGGRITGSDYDK 369
QY 357 RLITPGLFDHYVTLDMKLAWEPPGPIILPIIRVKADEAVAIANKSDPGLQSSVFTDFOK 416
DB 370 NRYKPTLVDPVPAKAKVSKETEGPIILPIITREDEALEMANDEFGLASVFTKDLAR 429
QY 417 AFDIANKLEVGTVHINNTGRCQDPNPFPLGLKSGAGVQGIKRYSIEMTNVKSIVL 472
DB 430 VEKVAALLEYGVAGNEIALSNPFT-FPGVGHSGFGRENGHYGMEBYIQVRFINL 484

RESULT 61

ID 08FUE1 PRELIMINARY; PRT; 499 AA.
AC 08FUE1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative aldehyde dehydrogenase.
GN OrderedlocusNames=CE0079;
OS *Corynebacterium efficiens*.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; *Corynebacterium*.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=YS-314;
RC MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of *Corynebacterium*
RT *efficiens*.";
RL Genome Res. 13:1572-1579(2003).
DR EMBL; AP005214; BAC16889.1; -.
DR HSSP; P56533; 1A45.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd. 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.
DR Complete proteome.
SQ SEQUENCE 499 AA; 53061 MW; F642B86E33C1BA8F CRC64;

Query Match 30.3%; Score 725; DB 2; Length 499;
Best Local Similarity 35.4%; Pred. No. 2.8e-35;
Matches 166; Conservative 102; Mismatches 185; Indels 16; Gaps 8;
QY 6 QNYVNGEM--KSSVQIIEILSPIDSSLGFPAMTRREVDHAMKAGREALPAMALTYVE 63
DB 22 QIFING-WVDGISESTTITNPFNGEBAVATIRMSQEDVDYAKATATAAVQEMAAVPEVK 80

QY 64 RAQYHKAADITIERDKEBIATVLAKETI-SKAYNAVTEVTRADILIRVAAEEGIRLSTSA 122
Db 81 KARIINKAAQVIEDRDALVELIRRESGSTALKANIEGLAMASL-REAAETPARIT--- 136
QY 123 DEGGHMDASTGHKLAVIRROPVGIYALAPVNVVNLSSGSKIALPALIGNVVMFKPPQOG 182
Db 137 --GQLPSTPEKSNVYFREFPLGVGVISPMNPPALSKRSVAPALACGNVGLVPASDT 194
QY 183 S-VSGIYVLAKAFARAGLPAGVFNTITGRGSEIGDYIVEHEVNFINTGSTPVGORIGL 241
Db 195 PIVGSTLLAHIFEAAGLPTGGLFNVVAGSEIGDFVEHRIIRLLISFTGSTVGGQVAT 254
QY 242 AG---MRPIMLECGKDAQIVLADLDNNAKQIVAGAYDVSQRCIAIKVLYVEEVA 297
Db 255 ATGKYIKKVALLELGGNAPMVVLLSPADINKAVAAASMSGSLFHQGOICMAVNRITVEAPY 314
QY 298 DELAKISEBNVAKLSVGPDPNAT--VTPYIDNSADFIESLVVDARQKAKELNEFKDG 356
Db 315 DQFLGFAARVSKVAVGQDLDATVGPVINDSQVESVSGKIDLABEQARAVVSGPIEG 374
QY 357 RLITFGLFDHVTLDMLKMAEERPGPILPIIRVKDAEAVAILANKSDPGLQSSVFTRPFOK 416
Db 375 RVIAHVADVTPEDELFRBELFGVVGKADNEHEALIELANDPREGLSGAVFTEDLOR 434
QY 417 AFDINKLEEVGTVINNTKGRGPDNFPPLGLKSGAGVQGIIRYSIEAMT 465
Db 435 GINFARVBEAGMTINDITVNDIAHVMFGEGKNSGIGRENGEMALDEFT 483

RESULT 62

QY 073EK2 PRELIMINARY; PRT; 483 AA.
AC 073EK2;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Succinate-semialdehyde dehydrogenase (NADP+) (EC 1.2.1.16).
GN Name=gabd; OrderedlocusNames=BCE0356;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=222523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Raeko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angluoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
adaptations and a large plasmid related to Bacillus anthracis pXOI,"
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
DR EMBL; AB01245; AAS39292.1; -
DR HSSP; P20000; IAG8.
DR TIGR; BCE0356; -
DR GO; GO:0004028; F:aldehyde dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0009013; F:succinate-semialdehyde dehydrogenase [NAD(P)-]; IEA.
DR GO; GO:0008153; P:metabolism; IEA.
DR InterPro; IPR02086; Aldehyd_dehydrotog.
DR InterPro; IPR010102; SSADH.
DR Pfam; PF00171; Aldehyd; 1.
DR TIGRFAMs; TIGR01780; SSADH; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
KM Complete proteome; Oxidoreductase
SQ SEQUENCE 483 AA; 52276 MW; AIDBFA7E035D42E6 CR664;

Query Match 30.2%; Score 722.5; DB 2; Length 483;
Best Local Similarity 35.5%; Pred. No. 3.7e-35;
Matches 170; Conservative 84; Mismatches 202; Indels 23; Gaps 8;

QY 8 YNNGWKSSVNOIELSPDSSLGFPVPMTRBEDHANKAGREALPMAALITVTRACY 67
Db 16 YINGEMITLQEOIEVNNPATKEIFATVPKGVTEAKQAVDAHAHEAFKWSKLTADRAVK 75

QY 68 LHKADITIERDKEBIATVLAKETISKAYNAVTEVTRADILIRVAAEEGIRLSTSADEGK 127
Db 76 LKNMPTLIDENKEEIALIMTEGQKPAEALGEVNVVANSFVEWVAEBSKVY-----YEM 130
QY 128 MDASTGHKLAVIRROPVGIYALAPVNVVNLSSGSKIALPALIGNVVMFKPPQOGSUGL 187
Db 131 IPASHPNRRIIVMKNOPGVMAAITPWNFPAAMIIRKAPALAACTAVKPKASOTPLPAL 190
QY 188 VLAARFAGLPAGVFNTITGRGSEIGDYIVEHEVNFINTGSTPVGORIGLAC--MR 245
Db 191 KLAELAHBADIPKGINIVTGSAPALADTWEDGRVKRVSFTSGTEIGKEIMASHAQTMK 250
QY 246 PIMLECGKDAQIVLADLDNNAKQIVAGAYDVSQRCIAIKVLYVEEVADEIAEKIS 305
Db 251 KVSLELGGHAPFIWMNDADLDKVAEAVIGSKFRNAGOCICTNNVFGVEEYVFAEAFQ 310
QY 306 ENVAKLSVGGPP-DNATVTPYIDNSADFIESLVVDARQKAKELNEFKR---DGLILT 360
Db 311 KSVGQLKVDGDFGDTVGLPIDRNAVSKVQEHADAIQKGTLYYGQKYAELDGHFMQ 370
QY 361 P---GLPHTVTLDMKMAEERPGPILPIIRVKDAEAVAILANKSDPGLQSSVFTRPFOK 417
Db 371 PTVGLADNTMLCMN---EETFGVAPVAKKTYEETERNANTPYGLAAVIFKDISQA 427
QY 418 FDIANKLEEVGTVINNTKGRGPD--NPPFLGLKSGAGVQGIIRYSIEAMTVKSIIVDM 474
Db 428 FOISEALEYGIIGLNDGL---PSVAQAPFGGFKSGIGREGGHHGIEEYLEIKTISLGL 483

RESULT 63

QY 081ZE2 PRELIMINARY; PRT; 483 AA.
AC 081ZE2; 061488; 06KXZ4;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Succinate-semialdehyde dehydrogenase (NADP+).
GN Name=gabd; OrderedlocusNames=BA0327; BAA0312; GBA00327;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RX MEDLIN=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eissen J.A., Gill S.R.,
RA Holtzapple E.K., Oekstad O.A., Helgason E., Ristone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plant R.D., Wolf A.M., Watkins K.L., Nietman W.C.,
RA Hazen A., Cline R.T., Redmond C., Twaite J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
closely related Bacteria,"
RN Nature 423:81-86(2003).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Raeko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Frazer C.M.;
RT "Bacillus anthracis comparative genomics,"
RL Submitted (May-2004) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sterner;
RA Bretin T.S., Bruce D., Chailcombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longshore J., Lucas S., Okinaka R.,

RA Richardson P., Rubin E., Tice H.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
 DR EMBL; AE017025; AAP24360.1; -
 DR EMBL; AE017334; AAT29418.2; -
 DR EMBL; AE017225; AAT52643.1; -
 DR HSSP; P51977; 1BX5.
 DR TIGR; BA0327; 1BX5.
 DR TIGR; GBA0327; -
 DR GO; GO:0004028; F:aldehyde dehydrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldehyd_dehydrog.
 DR InterPro; IPR010102; SSADH.
 DR Pfam; PF00171; Aldehyd; 1.
 DR TIGRfam; TIGR01780; SSADH; 1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
 DR Complete proteome; Oxidoreductase.
 SQ SEQUENCE 483 AA; 52258 MW; 04C6B3D3F648F277 CRC64;

Query Match 30.2%; Score 721.5; DB 2; Length 483;
 Best Local Similarity 35.7%; Pred. No. 4.3e-35;
 Matches 171; Conservative 82; Mismatches 203; Indels 23; Gaps 8;

QY 8 YVNGEKSSVNOIEILSPIDSSIGFVPAMTREEDVHAKAGREALPAMAALTVERAQQ 67
 Db 16 YINGEWITTOEQEIVNNPATKEIFATVPKCGVTEAKQVDAHAHFAFKSKLTADRAAK 75
 QY 68 LHKADIIERDKEEIAITVLAKETSKAYNASVTEVVTADLIRYAAEEGIRLSTSADEGK 127
 Db 76 LKKWFLIDENKEEIAIMTKEGKPPAALGEVYANVSFVEVYAEGRV-----YGEH 130
 QY 128 MDASTGHKLAVIRQVGVILAIAPYNYPNVLSGSKIAPALIGANNVMPKPTQSVSG 187
 Db 131 IPASHPNKRIILVKQVGVAAITPWNFPAMITRKVAPALAGCTAVVAPASQTPLTAL 190
 QY 188 VLAKAPAEAGLPAGVNTTIGRSEIGDYIVHEEVNFINTGSTPVGQIGLAG--NR 245
 Db 191 KLAEHLAHEADIPRGVINITYGSAKALADYMEDEGRKAKVSFTSGTEIGKELMASAQTMK 250
 QY 246 PIMLEIGKGAGIVLADADLDNNAKOIVAGAYDSGORTAIKRVILVEEVADELAEKIS 305
 Db 251 KVSLEIGGHAFFIVMDADLDKAVEAVISGKFRNAGQTCICTRVPVQEEVYEAFAVEKE 310
 QY 306 ENVAKLSVGDPP-DNATVTPVIDNSADFIESLVDAARQKAKELNEFRK----DGRLLT 360
 Db 311 KAVGQKLKVDGFGDGTTPVPLIDENAVSKVQEHIBDAIQGGTVLYGQGVAAELDGHFMQ 370
 QY 361 P---GLFDHTVTLDMKLAMEPFGPIPIIRVDAEEVAVALANKSDPGLQSSVFTTRDPQKA 417
 Db 371 PTVIGLANDTMLCMN---EETFGPVAPVAKFTVEEYIERAHTPPGLAAVIFTKDISQA 427
 QY 418 FDIANKLEVGTVHINKTKGRPD--NPPLGLKSGAGVQGIYRSIEAMTNVKSIVLDM 474
 Db 428 FOISEALEYGIIGLNDGL---PSVAQAPFGGFKESGIGREGGHGIEEYLEIKYISLGL 483

RESULT 64

Q6HP75 PRELIMINARY; PRT; 483 AA.
 ID O6HP75
 AC O6HP75
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DN Succinate-semialdehyde dehydrogenase (NAD(P)+) (EC 1.2.1.16).
 OS Name-gad; Ordered;cuenNames=BT9727.0295;
 OS Bacillus thuringiensis (subsp. konkukian).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=180856;
 OX NCB1
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=97-27;
 RT Bretzin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hall K.,

RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Tice H.;
 RL "Complete genome sequence of Bacillus thuringiensis 97-27."
 CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
 DR EMBL; AE017355; AAT61348.1; -
 DR GO; GO:0004028; F:aldehyde dehydrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0009013; P:succinate-semialdehyde dehydrogenase [NAD(P) .]; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldehyd_dehydrog.
 DR InterPro; IPR010102; SSADH.
 DR Pfam; PF00171; Aldehyd; 1.
 DR TIGRfam; TIGR01780; SSADH; 1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
 DR Complete proteome; Oxidoreductase.
 SQ SEQUENCE 483 AA; 52258 MW; 04C6B3D3F648F277 CRC64;

Query Match 30.2%; Score 721.5; DB 2; Length 483;
 Best Local Similarity 35.7%; Pred. No. 4.3e-35;
 Matches 171; Conservative 82; Mismatches 203; Indels 23; Gaps 8;

QY 8 YVNGEKSSVNOIEILSPIDSSIGFVPAMTREEDVHAKAGREALPAMAALTVERAQQ 67
 Db 16 YINGEWITTOEQEIVNNPATKEIFATVPKCGVTEAKQVDAHAHFAFKSKLTADRAAK 75
 QY 68 LHKADIIERDKEEIAITVLAKETSKAYNASVTEVVTADLIRYAAEEGIRLSTSADEGK 127
 Db 76 LKKWFLIDENKEEIAIMTKEGKPPAALGEVYANVSFVEVYAEGRV-----YGEH 130
 QY 128 MDASTGHKLAVIRQVGVILAIAPYNYPNVLSGSKIAPALIGANNVMPKPTQSVSG 187
 Db 131 IPASHPNKRIILVKQVGVAAITPWNFPAMITRKVAPALAGCTAVVAPASQTPLTAL 190
 QY 188 VLAKAPAEAGLPAGVNTTIGRSEIGDYIVHEEVNFINTGSTPVGQIGLAG--NR 245
 Db 191 KLAEHLAHEADIPRGVINITYGSAKALADYMEDEGRKAKVSFTSGTEIGKELMASAQTMK 250
 QY 246 PIMLEIGKGAGIVLADADLDNNAKOIVAGAYDSGORTAIKRVILVEEVADELAEKIS 305
 Db 251 KVSLEIGGHAFFIVMDADLDKAVEAVISGKFRNAGQTCICTRVPVQEEVYEAFAVEKE 310
 QY 306 ENVAKLSVGDPP-DNATVTPVIDNSADFIESLVDAARQKAKELNEFRK----DGRLLT 360
 Db 311 KAVGQKLKVDGFGDGTTPVPLIDENAVSKVQEHIBDAIQGGTVLYGQGVAAELDGHFMQ 370
 QY 361 P---GLFDHTVTLDMKLAMEPFGPIPIIRVDAEEVAVALANKSDPGLQSSVFTTRDPQKA 417
 Db 371 PTVIGLANDTMLCMN---EETFGPVAPVAKFTVEEYIERAHTPPGLAAVIFTKDISQA 427
 QY 418 FDIANKLEVGTVHINKTKGRPD--NPPLGLKSGAGVQGIYRSIEAMTNVKSIVLDM 474
 Db 428 FOISEALEYGIIGLNDGL---PSVAQAPFGGFKESGIGREGGHGIEEYLEIKYISLGL 483

RESULT 65

DHAL BACST STANDARD; PRT; 488 AA.
 ID P42329;
 AC P42329;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Aldehyde dehydrogenase, thermostable (EC 1.2.1.3).
 GN Name-alDHT;
 OS Bacillus stearothermophilus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 NCBI_TaxID=1422;
 OX NCB1
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SIC1;
 RA Imanaka T., Ohta T., Sakoda H., Widyastuti N., Matsunaka M.,
 RT "Cloning, nucleotide sequence, and efficient expression of the gene coding for thermostable aldehyde dehydrogenase from Bacillus

RT steaerothermophilus, and characterization of the enzyme.";
RL J. ferment. Bioeng. 76:161-167(1993).
CC -!- FUNCTION: Oxidizes several aliphatic aldehydes, particularly c6-
CC aliphatic aldehyde and hexanal, but do not oxidize benzaldehyde.
CC -!- CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
CC -!- COFACTOR: Requires either NAD or NADP as a coenzyme.
CC -!- PATHWAY: Ethanol utilization; second step.
CC -!- SUBUNIT: Homotrimer.
CC -!- MISCELLANEOUS: The optimum temperature is around 55 and 60 degrees
CC Celsius.
CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-slb.ch/announce/>
CC or send an email to license@1sb-slb.ch).
CC -----
CC EMBL: D13846; BAA02975.1; -
CC PIR: I39769; I39769.
CC DR HSSP: P51977; 1BXS.
CC DR InterPro: IPR002086; Aldehyde_dehydr.
CC DR Pfam: PF00171; Aldehyd; 1.
CC DR TIGRFAMs: TIGR01804; BADH; 1.
CC DR TIGRFAMs: TIGR01722; MMSDH; 1.
CC DR TIGRFAMs: TIGR01780; SSADH; 1.
CC DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
CC DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; 1.
CC DR NAD: Oxidoreductase.
CC FT NP BIND 233 NAD (ADP part) (By similarity).
CC FT ACT_SITE 255 255 By similarity.
CC FT ACT_SITE 289 289 By similarity.
CC SQ SEQUENCE 488 AA; 52915 MW; 29824451965095D CRC64;

Query Match 30.2%; Score 721.5; DB 1; Length 488;
Best Local Similarity 34.4%; Pred. No. 4.3e-35;
Matches 166; Conservative 104; Mismatches 192; Indels 21; Gaps 9;

QY 3 KEYQYVNGEMKSSV-NQIE-ILSPIDSSLG-FVPAMTREEDVHAMKAGREALPMAAL 59
DB 8 KTYFYVINGNWSSVSNVNPSPINPANHDIYGVYRSTLEEDVNEAVTANERQTSWMKR 67
QY 60 TYVERAOYHRAADIIEERKEIATVLAKEISKAVNASTVEVTRADLIRYAAEGIRLS 119
DB 68 SGVEREEYLYKAAHILEQCLQDIATMTREMGKTLAEAKAETVRGVHILRYAAGEGAR-- 125
QY 120 TSADEGKMDASTGHTKLAVIRROVGIYLAIPYNYVNLGSKIAIPALIGNVVFKFP 179
DB 126 ---KIGDVI PSSDSBELFTTRVPLGVGVISIPMNFPVAIPYKMAPALVGNVTLKPA 182
QY 180 TGGSVGLVLAFAFAAGLPAGVFNITTRGSEIGDYIVHEBEVNFINTGSTPVGORIG 239
DB 183 SETAVTAAKVIICFHEAGPRKGVNNVCGSGSVVCGGIANHPRIDVFTFGSTVVGKVG 242
QY 240 KLAQMR--PIMLEIGKDAIGVLADADLNAKQIYAGAVDYSQGCRTAKRVLVBEBA 297
DB 243 RAAFERGAATYOLEMGSKNPVIYAKDADLDAVEGTISGGRSTGCKTATSRPFIERYEV 302
QY 298 DELAEKISENVAKLTVGDPFDNAT-VTPYIDNSADPFIESLVVDARQKAKELNEFR-- 354
DB 303 EEPFKALKLBRVQKLKINGLDAETWNGPCASBSQFHTVLSYIEKGSBQAKLIYGNRCL 362
QY 365 -----DGRLLTGLPDHVTLDKMLAMBEPGPILPIIRVKADEAVAIANKSDPGLQSSV 409
DB 363 ESELANGFVEBPITFEVDVLDQMTIAREEPIFGPVTLAIQVDSIEBAIKLANDTEYGSASI 422
QY 410 FTTRDFKADIANKLELVGVTHINNKTRGRDNPFPGLK--GSGAIVGQIRYSIEAMTV 467
DB 423 YTKNIGNALEFTKIDIEAGLIKVAETAGVFPQAPFGMKOSSSHSREG--QAATIEFTSI 481
QY 468 KSI 470

DB 482 KTV 484
|:|
RESULT 66
Q92DA3 PRELIMINARY; PRT; 488 AA.
ID Q92DA3
AC Q92DA3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lin0913 protein.
GN OrderedLocustNames=lin0913;
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_Taxid=1642;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaeser P., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Chardit A., Chetouani F., Couve E., de Darvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunz F., Kurupkat G.,
RA Madeno E., Maitouram A., Mata Vicente J., Ng E., Nedjati H.,
RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT "Comparative genomics of Listeria species";
RL Science 294:849-852(2001).
DR EMBL: AL596166; CAC96145.1; -.
DR PIR: A11546; A11546.
DR HSSP: P05091; 1004.
DR ListerList; LIN0913; -.
DR GO: GO:0016491; P:oxidoreductase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002086; Aldehyd_dehydrog.
DR Pfam: PF00171; Aldehyd; 1.
DR Complete Proteome.
SQ SEQUENCE 488 AA; 53351 MW; 3270FF2486CB08CE CRC64;

Query Match 30.1%; Score 720; DB 2; Length 488;
Best Local Similarity 33.6%; Pred. No. 5.3e-35;
Matches 160; Conservative 100; Mismatches 198; Indels 18; Gaps 6;

QY 8 YVNGEMKSSVNO--IEILSPIDSSLG-FVPAMTREEDVHAMKAGREALPMAALTVYERA 65
DB 16 PINGMKINGDNKERTDYINPANGVIYAKIQAQPRETQALKAADAPDMAKMLARV 75
QY 66 QYHRAADIIEERKEIATVLAKEISKAVNASTVEVTRADLIRYAAEGIRLSADEG 125
DB 76 KLHNIADIMBERADTLAKIMTLEQCKPLSKSGEVLTCAGNFRPAABEARL-----YG 130
QY 126 GMDASTGHTKLAVIRROVGIYLAIPYNYVNLGSKIAIPALIGNVVFKFPPTQSSVS 185
DB 131 ETTPANNHAF-IVKQPIGVAAITPWNFGQWTRKLAPLATGNTVTLKPSGDTPLS 189
QY 186 GLVLAFAFAAGLPAGVFNITTRGSEIGDYIVHEBEVNFINTGSTPVGORIGTLAAG-- 243
DB 190 ALAIFRIFREAGLPKGVANIVWGSKEIGFTLTSDVDRKLTFTOSTVVGOTLPFGQSDT 249
QY 244 MRPIMLEIGKDAIGVLADADLNAKQIYAGAVDYSQGCRTAKRVLVBEBADELAEK 303
DB 250 LKKSILEIFGHAAPFIVFADANIEAAVDDLITAKFRNNGQVCVSPRIRIVAAEIKKEFTKT 309
QY 304 ISENAKLSVGDPPFNATVTPYIDNSADPFIESLVVDARQKAKELNEFR-----DG 356
DB 310 IVAKTEKIKVNGGLDGVNVGPIREDADIKDKQIKNATDKGAKVLTGSEIRLTGSDYDKG 369
QY 367 RLITPGLPDHVTLDKMLAMBEPGPILPIIRVKADEAVAIANKSDPGLQSSVFTTRDFOK 416

Db 370 NFKPVLVDNVTBMDIFYETFGPVLPILVFENEDAIEMADSEGLSYITTKDLAR 429
QY 417 AFDIANKLEVGTVHINNKTGRGPDNFPFLGLKSGAGVGGIRYSIEMTNVKSIVL 472
Db 430 VEKVGAALEGVGVGANEIAISNPET--PFGVGKHSFGFRENHGHGMEEYIQVKEFINL 484

RESULT 67
Q63GQ3 PRELIMINARY; PRT; 483 AA.
ID 063GQ3
AC 063GQ3; 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Succlinate-semialdehyde dehydrogenase (NAD(P)+) (EC 1.2.1.16).
GN Name=gadD; ORFNames=BTZK0299;
OS Bacillus cereus ZK.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=288681;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZK;
RA Bretzin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus cereus ZK";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000001; AAV19940.1; -.
KM Oxidoreductase.
SQ SEQUENCE 483 AA; 52243 MW; 30C61FA31420849 CRC64;

Query Match 30.0%; Score 718.5; DB 2; Length 483;
Best Local Similarity 35.5%; Pred. No. 6.5e-35;
Matches 170; Conservative 83; Mismatches 203; Indels 23; Gaps 8;

QY 8 YVNGEKSSVNOQIEILSPIDSSLGVPANTREEDVHAMKAGREALPAMALTYVERAQY 67
Db 16 YINGEWITLQEOIEVNNPATKEIFATVPKGVTEAKQAVDAHAFAKSWSKLTRAADRAAK 75
QY 68 LHKRAADIIBDKKEIATVLAKEISKAYNASVTEVVRADILRYAAERGRILSTSADEGK 127
Db 76 LKMKFPLIDENKEEILAIIMTKEQKPEALGENTANSFVEMYABEGKKV----YGEW 130
QY 128 MDASTGHLAVIRROPVGIYLAIPYVPVNLGSKIAPIALIGNVMMFEPPTQSGVSL 187
Db 131 IPASHNKRILVMKOPGVMAALTTPNMFPAAMITRKVAPALAGCTAAVVPASQTPILAL 190
QY 188 VLAKAFPAEAGLPAGVFNITGRGSEIGDYVEHEEVNFINTGSTPVGQRIGKLAG--MR 245
Db 191 KLAEILHAEADIPKGVINIVTGSAAKAIADTMEDGRVAKVSFTGSTELGKELMASAATMK 250
QY 246 PIMLELGKQAGIVLADADIDNNAKOIVAGAYVYSGORCTAIRVILVVEVDELAELAKS 305
Db 251 KVSLELGGAHPFTVMNDADIDKAVEATVSGKFRNAGTCTCTNRVFOEVEVYEAFAVEKQ 310
QY 306 ENYAKLSVGDPE--DNATVTPVIDNSADFIESLVVDARQKAEINLEFKR----DGRILTL 360
Db 311 KAVGOLKVGDFGDFGDTTVPILIDENAVSKVQEHIEBALQKGVIVLGGQVVALDGHFMG 370
QY 361 P---GLPDHVTLDMLKAMEEPGPILPIIRVKDAEBAVAIAKNSDFGLQSSVFTDPQK 417
Db 371 PTYIGLANDTMLCMN--EETFGPVAVAKFTVDEYIERANHTPYGLAAVYIFTKDISQ 427
QY 418 FDIANKLEVGTVHINNKTGRGPD--NPPFLGLKSGAGVGGIRYSIEMTNVKSIVLDM 474
Db 428 FOISEALEYIGLIDGL--PSVAQAPFGGFYBSGIGREGHFGIBEYIEIKYISLGL 483

RESULT 68
Q8Y819 PRELIMINARY; PRT; 488 AA.
ID 08Y819
AC 08Y819;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lmo0913 protein.
OS OrderedLocustNames=lmo0913;
OC Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaeser P., Frangeul L., Buchrieser C., Ruenliok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Cheoutani F., Couve E., de Daruvar A., Delhoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
RA Eutlier K.-D., Fehli W., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapatk G.,
RA Madueno E., Maltournam A., Mata Vicente J., Ng E., Nedjati H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tietzer A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT "Comparative genomics of Listeria species";
RL Science 294:849-852(2001).
DR EMBL; AL591977; CAC98991.1; -.
DR PIR; A11188; A11188.
DR HSSP; P05091; 1004.
DR LIPIDList; LMO0913; -.
DR GO; GO:0016491; Fmocidoreductase activity; IEA.
DR GO; GO:0008152; Pimicetabolism; IEA.
DR Pfam; PF00171; Aldehdh; 1.
KM Complete proteome.
SQ SEQUENCE 488 AA; 53166 MW; 8EA0A920B1CE02A2 CRC64;

Query Match 30.0%; Score 718; DB 2; Length 488;
Best Local Similarity 33.8%; Pred. No. 7e-35;
Matches 161; Conservative 96; Mismatches 201; Indels 18; Gaps 6;

QY 8 YVNGEKSSVNO--IEILSPIDSSLGVPANTREEDVHAMKAGREALPAMALTYVERA 65
Db 16 FINGKMTDGDNKETKDIYVNPANGDYIAKIQAGPSETKAIIRAKADAFDMAKMELADYV 75
QY 66 QYLKRAADIIBDKKEIATVLAKEISKAYNASVTEVVRADILRYAAERGRILSTSADEG 125
Db 76 KILHKIADIMBKADTLAKIMTLBOGKPLKESKGEVLTEVENFRFAEABARL-----YG 130
QY 126 GKMDASTGHLAVIRROPVGIYLAIPYVPVNLGSKIAPIALIGNVMMFEPPTQSGVS 185
Db 131 ETIPAPNNHAF--IVKKOPIGVVAALTTPNMFPGMWYRKLAIPALATGNTIVLKPSGDTPLS 189
QY 186 GLYLAKAFPAEAGLPAGVFNITGRGSEIGDYVEHEEVNFINTGSTPVGQRIGKLAG-- 243
Db 190 ALAIFEIFEEAGLPAGVANIVTGSSEKEIGETLTSDDVAKLTFSTGKVGQTLFKQSADT 249
QY 244 MRPIMLELGKQAGIVLADADIDNNAKOIVAGAYVYSGORCTAIRVILVVEVDELAELAK 303
Db 250 LKKSILELGGAHPFIYFDANIDAAVNDVAAKFPNNGVCVSPNRIFAKKEKEFTYA 309
QY 304 ISENVAKLSVGPEDFNATVTPYVIDNSADFIESLVVDARQKAEINLEFKR-----DG 356
Db 310 LVAKEVQLKVGNGLDVNVNGLPIREPAIDIKOKUKNAEKAQKAVLTGGGRITGSDYDKG 369
QY 357 RLTLPGFPHVTLDMLKAMEEPGPILPIIRVKDAEBAVAIAKNSDFGLQSSVFTDPQK 416
Db 370 NFKPVLVDNVTBMDIFYETFGPVLPILVFENEDAIEMADSEGLSYITTKDLAR 429
QY 417 AFDIANKLEVGTVHINNKTGRGPDNFPFLGLKSGAGVGGIRYSIEMTNVKSIVL 472
Db 430 VEKVGAALEGVGVGANEIAISNPET--PFGVGKHSFGFRENHGHGMEEYIQVKEFINL 484

RESULT 69

```
Q8X950
ID Q8X950 PRELIMINARY; PRT; 482 AA.
AC Q8X950; Q7ABE9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Succinate-semialdehyde dehydrogenase, NADP-dependent activity.
GN Name=gabd; OrderedlocusNames=EG63522, z3959;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_TaxID=83334;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 050952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakiwa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
RN
CC -I- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
DR EMBL; AB005495; AAG57768.1; -.
DR EMBL; AP002562; BAB36945.1; -.
DR PIR; B91069; B91069.
DR PIR; D85913; D85913.
DR HSSP; P20000; 1A68.
DR HSSP; Q28399; 1O9J.
DR GO; GO:0004028; F:aldehyde dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR InterPro; IPR010102; SSADH.
DR Pfam; PF00171; Aldehyd; 1.
DR TIGRfam; TIGR01780; SSADH; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Oxidoreductase; Complete proteome.
SQ
SEQUENCE 482 AA; 51817 MW; 0CF3D4790A9EAD15 CRC64;

Query Match 30.0%; Score 717; DB 2; Length 482;
Best Local Similarity 33.9%; Pred. No. 8e-35;
Matches 162; Conservative 92; Mismatches 208; Indels 16; Gaps 7;

QY 6 QYVNGEMKSSVNV-OIEILSPIDDSLGFPVPMTEBEVDHAKAGREALPMAALITVYE 63
DB 12 QALINEMEDANNNGEVIDVTPNANGDKLSVPRMGADETPAIDAANRALPVPFALTAKE 71
QY 64 RAQYLKAAADIIRDEKEIATVLAKEISKAYNASVTEVVRTADLIRYAABEGRLSTASD 123
DB 72 RANILNEMEMEHODDLARLMTLBQGRFLAEAKGEISYASFIEMFAEBGKRTYGDIT 131
QY 124 EGGKMDASTGHKLAVIRQPVGIVLAIPYNYEVNLSGSKIAPALIGGVNMFKPPTGS 183
DB 132 PGHQAD-----KRLIYIKQIGVTAITPNNFPAAMITRKAGPALAAGCTMVLKPSQTP 186
QY 184 VSGVLTAFAFAEAGLPAQVFNITTGSGSEIGDIYIEHEEVNFNFPGSTPVGQRIGKL 243
DB 187 FSLALAEALIRAGIGPAGVFNVTGSAAGVAGNELTSPNPLVRKLSFTGSGTEIGRLMEQCA 246
```

```
QY 244 --MRPIMELGKGADGIADADLDNAKOIVAGAYDYSGORCTAIRKLVVEEVADELA 301
DB 247 KDIKVYSLELGNNAFFIYFDADLDKAVEGALASFRNAGGTCVCAANLYVQDDGYDBFA 306
QY 302 EKISENVAKLSVGDPFD-NATVTPIYDINSADFIESLVVDARQKAGEL---NEFKPDGR 357
DB 307 EKLQQAASKLHIGDLDGVTIGPLIDEKAVAKBEHIALDLBKGARVVCCKAHERGQN 366
QY 358 LITPLPFIHYTLDMKLAMEEPFGPLPIIRYKDAEENVAINKSPFGIQQSVFTRPQKA 417
DB 367 FFGPITLVDPANAASVEETFGPLAPLFRKDEADVTAQANDPEFGLAATFYADLSRV 426
QY 418 FDIANKLEVTGVIHINKTG-RGPDNFPFLGKSGAGVQGIYSIEAMTVNSIYLDM 474
DB 427 FRVGALEHYGVGIN--TGIISNEVAPFGIKASGLGSGSKYGIEDYLEIKYMGIGL 482

RESULT 70
Q9HK01
ID Q9HK01 PRELIMINARY; PRT; 493 AA.
AC Q9HK01;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Probable aldehyde dehydrogenase.
GN OrderedlocusNames=Ts0809;
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmales;
OC Thermoplasmataceae; Thermoplasma.
NCBI_TaxID=2303;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001; DOI=10.1038/35035069;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Fritsman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513 (2000).
RN
CC -I- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
DR EMBL; AL445065; CAC11938.1; -.
DR HSSP; Q28399; 1O9J.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ
SEQUENCE 493 AA; 54782 MW; 2F6A79DD96BETFF5 CRC64;

Query Match 29.9%; Score 716.5; DB 2; Length 493;
Best Local Similarity 34.6%; Pred. No. 8.e-35;
Matches 167; Conservative 97; Mismatches 191; Indels 27; Gaps 9;

QY 8 YNNGEM--KSSVNOEILSPIDDSLGFPVPMTEBEVDHAKAGREALPMAALITVYRA 65
DB 6 YIDGPMVNSSSGKTYDKSPVYGVYGRFEATIRVDVRAIDAADATAMANDLGSVERS 65
QY 66 QYLAKRAADIIRDEKEIATVLAKEISKAYNASVTEVVRTADLIRYAABEGRLSTADEG 125
DB 66 KIYIAKLEIENRARELNIEMENKGVKEAKEVEDVDIVQIYVAAWMARKLNGEVEVG 125
QY 126 GMDASTGHKLAVIRQPVGIVLAIPYNYEVNLSGSKIAPALIGGVNMFKP--PTGS 183
DB 126 -----TSGHRRKIFYQVYGVYIVALTIPNPFAGVARKIAPALIGTVVVKPSSDPTGS 180
QY 184 VSGVLTAFAFAEAGLPAQVFNITTGSGSEIGDIYIEHEEVNFNFPGSTPVGQRIGK--L 241
DB 181 AEWY--RFEVYAGPKGVLFNITGRGSEIGDIYIEHKKVNLITMTGSLTAAGQRIMQAS 238
QY 242 AGMRPIMELGKGADGIADADLDNAKOIVAGAYDYSGORCTAIRKLVVEEVADELA 301
```



```

Db      239 ANNAKILIELGCGAPPMWMDADMDALKTLMMAKMMNGOSCTIAERLYVHEDIYDTFM 298
      302 EKISENVAKLISVGDPPDNATVTPEVDNSADFIESLVVDARQKAK-----ELINEK 353
      299 SFPEVLSRKLALGDP-KNADMGPLINKGALQATSEIVEEKESGAKILFGSGQPSISGPIY 357
      354 RGRRLTLPGLFVDVTLDMKLAMEEPFGPIPIIRVDAEBAVALANKSPFGLQSSVPTPD 413
      358 RUGYFPLFTTIGADOKSKIPOEIRPAVYGARKISSVEEMVYDLADSKVGLASVLTPTXD 417
      414 FOKAPFANKLEVGTHINNKTRGPDNFP--FLGKSGAGVQGIIRYSIEAMTVKSIYL 471
      418 PNIFPASERIRGELIYVMP--GPEAGOGHTGFRMTGQAGSGSKTGISYLKKNILY 474
      472 LD 473
      475 VD 476

```

RESULT 71

```

ID      081DV8      PRELIMINARY;      PRT;      474 AA.
AC      081DV8;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Succinate-semialdehyde dehydrogenase (EC 1.2.1.16).
GN      OrderedLocNames=BC2241;
OS      Bacillus cereus (strain ATCC 14579 / DSM 31).
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC      NCB1_TaxID=226900;
RN      NCB1_TaxID=226900;
RX      SEQUENCE FROM N.A.
RX      MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
RA      Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA      Kapral V., Bhattacharya A., Reznik G., Mikhailova N., Japich A.,
RA      Chu L., Mazur M., Goldsman E., Larsen N., D'Souza M., Walunas T.,
RA      Greckhik Y., Pusch G., Haselkorn R., Fongstein M., Ehrlich S.D.,
RA      Overbeek R., Kyriides N.C.;
RT      "Genome sequence of Bacillus cereus and comparative analysis with
RT      Bacillus anthracis.";
RL      Nature 423:87-91(2003).
DR      EMBL; AEO17005; AAP09206.1; -.
DR      HSSP; Q28399; 109J.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0009013; F:succinate-semialdehyde dehydrogenase [NMD(P. ., IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      InterPro; IPR002086; Aldehyd_dehydrog.
DR      Pfam; PF00171; Aldehyd; 1.
KW      Complete proteome; Oxidoreductase.
SQ      SEQUENCE 474 AA; 51868 MW; 598AE1E3B81D5C68 CRC64;

```

Query Match 29.9%; Score 716; DB 2; Length 474;

Best Local Similarity 35.1%; Pred. No. 9e-35;

Matches 165; Conservative 98; Mismatches 199; Indels 8; Gaps 5;

```

QY      8 YVNGEKSSVNOEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMAALTVERAY 67
      6 YVNGSKSVVTVTPPLAIFYSEETLAIAGSTEDVEKAVTAAKNAATKANKKSAVRAPI 65
      68 LHKAAIDIERDKKEIATVLAKESIKAYNASVTEVVRADLIRYAAEGIRLSTSADEGK 127
      66 LKKAQMDERREFEAIIKAKAPIRARGEVDRTVQYKFAAEAKRI--YGETLP 122
      128 MDASTG--KTLAIVRQPGVIVLAIPYNPVNLSSKTAIPALIGNVVMFKPPTGSGVS 185
      123 LDKAPAGDGIATITRKPIGIVICATIPNPPLNLVAHKGPALAAAGTVLKKADPTPIS 182
      186 GLVLAFAFAEAGLPAGVFNTITGRGSEIGDYIVEHEEVNFINTGSPVQRIQKLAGMR 245
      183 SYALVELFERAGLPNGALNLTISGPGTVGSAIVKNDVVASITTTGSPKVGIGIKQKAGLK 242

```

```

QY      246 PIMLEIGKADGIVLADADL-DNAKQIVAGAYDYGRCRTAKRVLVEEVADELEAKI 304
      243 RVTLEIGSNAVITIEDVEHLEITERVYKMGAFVANNQVCISQVRFVHETIHELSTG 302
      305 SENVAKLISVGD--PNNATVPYIDNSADFIESLVVDARQKAKEL-NBPKEDGRLLTNG 362
      303 KKAMSVVVGDPSLBETDVSALISKDVERIDMWQEAIKEGATVLCGKKQDARIFEPT 362
      363 LFDHTVLMKLAMEEPFGPIPIIRYKDAEEVAVALANKSPFGLQSSVPTPDQAFDIAN 422
      363 VLTNPDVHVSQCOEYVFGPLMTVNTFKERDEALIQVNNNSRYGLQAGVFNTNLFKMAIRD 422
      423 KLEVGTHINNKTRGPDNFPFLGLKSGAGVQGIIRYSIEAMTVKSIYL 472
      423 ELEVGVMINDI.PFRVDHMPYGVKESGTGREGIKTALTEEMTMLVCI 472

```

RESULT 72

```

ID      09HQ22      PRELIMINARY;      PRT;      480 AA.
AC      09HQ22;
DT      01-MAR-2001 (TrEMBLrel. 16, Created)
DT      01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Glyceraldehyde-3-phosphate dehydrogenase.
DE      NameGap; OrderedLocNames=VNG0937G;
OS      Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC      Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC      Halobacteriaceae; Halobacterium.
OC      NCB1_TaxID=64091;
RN      NCB1_TaxID=64091;
RX      SEQUENCE FROM N.A.
RX      MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA      Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA      Shukla H.D., Lasky S.R., Balliga N.S., Thorsen V., Shrogha J.,
RA      Swartzell S., Weir D., Hall J., Dahl T.A., Meli R., Goo Y.A.,
RA      Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA      Maddocks D.G., Jablonki P.B., Krebs M.P., Angevine C.M., Dale H.,
RA      Isebnarger T.A., Beck R.F., Pohlischer M., Spedlich J.L., Jung K.-H.,
RA      Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA      Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT      "Genome sequence of Halobacterium species NRC-1";
RL      Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR      EMBL; AEO05031; AAG19367.1; -.
DR      PIR; C84250; C84250.
DR      HSSP; O57693; 1KY8.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      InterPro; IPR002086; Aldehyd_dehydrog.
DR      Pfam; PF00171; Aldehyd; 1.
DR      PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
DR      PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.
KW      Complete proteome.
SQ      SEQUENCE 480 AA; 50483 MW; FB86939702082CD2 CRC64;

```

Query Match 29.8%; Score 713.5; DB 2; Length 480;

Best Local Similarity 35.0%; Pred. No. 1.3e-34;

Matches 165; Conservative 86; Mismatches 210; Indels 11; Gaps 6;

```

QY      8 YVNGEKSSVNOEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMAALTVERAY 67
      13 FYAGEVNSDGEPLPSDLADGGTFASVAADTSDAERLISAATGVNADLRETTVERVEM 72
      68 LHKAAIDIERDKKEIATVLAKESIKAYNASVTEVVRADLIRYAAEGIRLSTSADEGK 127
      73 LESIADGIRRRDELAELAVVRBAKRPISARBEVGSAAERFRAVGE-LRHITGEVPTG- 130
      128 MDASTGHR--LAVIRQPGVIVLAIPYNPVNLSSKTAIPALIGNVVMFKPPTGSGVS 185
      131 --TTGHEHMOAIVKHEPMTVLCITPPYPLSTALQVAPAIAGNAVIVPASTPTIS 188
      186 GLVLAFAFAEAGLPAGVFNTITGRGSEIGDYIVEHEEVNFINTGSPVQRIQKLAGMR 245

```


Db 189 GALLIADIADAGL.PDGAAYNFVPESSVIGDPLASDARVDIAIMTSSGAGEHVARQGIT 248
QY 246 PIMTELGGAGAGIYVADADLDNAKQIVAGAYDSGORCTAIKRYLVVVEVADELAEKIS 305
Db 249 RLHMLGGAAPVIEVEDADLDNAADATRGSLKTIKGORCSAVSRVLAHSEVDELVSRI 308
QY 306 ENVAKLVSVDPPD-NATYTPVIDNSADFIESLVVDARQKAKELNEFKR---DG-RLIT 360
Db 309 DAMAEMSIGDLFDPTTIGPLVSADQADWVAELVDAVDRGATVVRGGERHVEDGVHYE 368
QY 361 PGLPHVTLDMKLAEMEPGPPLIPIIRYVDABEVAIAIKSPFGQSSVFTFDFOKAFPI 420
Db 369 PFLADVPDARIIVDEBQGPVCAVTVTVDEDDAVRTANGSELALDAVFTADHDRAMRV 428
QY 421 ANKLEVGTVHINNKTRGPDNFPFLGLKSGAGVQGIIRYSIEAMTNVKSIVL 472
Db 429 AERVAGAVRINGAISHGLGVFPFGNDASGIGREGDSTHEFVREKSIIL 480
RESULT 73
057693 PRELIMINARY; PRT; 501 AA.
AC 057693:
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (NADP+) [EC 1.2.1.9].
GN Name=GANP;
OS Thermoproteus tenax.
OC Archaeae; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Thermoproteus.
OX NCBI_TaxID=2271;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KRA 1;
RA Brunner N.A.;
RT Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL, Y10625; CAA71651.1; -.
DR PIR, T44939; T44939.
DR PDB, 1KY8; X-ray; A=1-501.
DR PDB, 1UXN; X-ray; A=-.
DR PDB, 1UXP; X-ray; A=-.
DR PDB, 1UXO; X-ray; A=-.
DR PDB, 1UXR; X-ray; A=-.
DR PDB, 1UXT; X-ray; A=-.
DR PDB, 1UXV; X-ray; A=-.
DR PDB, 1UXW; X-ray; A=-.
DR GO, GO:000886; F:glyceraldehyde-3-phosphate dehydrogenase (N. .; IEA.
DR GO, GO:001649; F:oxidoreductase activity; IEA.
DR InterPro, IPR002086; Aldehyde_dehydrog.
DR Pfam, PF00171; Aldehyd_1.
DR Oxidoreductase.
KW SEQUENCE 501 AA; 54090 MW; 22F1C0084282EE1B CRC64;
Query Match 29.8%; Score 713; DB 2; Length 501;
Best Local Similarity 37.0%; Pred. No. 1.4e-34;
Matches 180; Conservative 83; Mismatches 196; Indels 28; Gaps 9;
QY 5 YONYNGEMKSVNIEILSPIDSSIGFVPAMTRREVDHAM----KAGREALPMAA-- 58
Db 19 YPSIAGEGGSGGQIEVSPIDLTIAVISPSSREVRRTLDVLFKGR-----WSAD 73
QY 59 LTVVERAQYLKRAADIIEERDKKEIATVLAKEISKAVNASVTEVVTADLIRVAAEGRL 118

Db 74 MPTGERLAVLKRADIIERNDVFAEVLWMNAKPRKSAVGEVRAADVRL-LAELOLKK 132
QY 119 STSADEGKMASTGCHKLAIVRRQPVGIYLAIPRYNVISGSKIALPAGNVWMEKP 178
Db 133 IGGDVIIPDWTYDTLETETGVRREPLGVAAITPKNVPLFPAVKNKITYSFTYGNAAVVKP 192
QY 179 PTQSGSLVLAKAFAEAGLPAGVFNITGSGSEIGDVIYHEEVNFINTGSPFVGQRI 238
Db 193 SISDPLPAAAVKALLDGFPDDAIALNLFGKE-AEKIVADDRVAVSTFGSTVEGRV 251
QY 239 GKLAGMRPIMTELGGAGIYVADADLDNAKQIVAGAYDSGORCTAIKRYLVVEEVD 298
Db 252 VKVGSGQYVMELGGDPAIYLEDADLDLADKIARGIYSVAGORDAIKLVLAERPYYG 311
QY 299 ELAEKISENVAKLSVGPDP-NATYTPVIDNSADFIESLVVDARQKAKELNEFKRGR 357
Db 312 KLVEEVAARLSSLRGDRDPTVDVGPLISSAEDENMAALIDAVEKGRVLT---AGGR 367
QY 358 LITPLGPDHVTL-----DMKLAEMEPGPPLIPIIRYVDABEVAIAIKSPFGLOQS 408
Db 368 RLGP-TYIQTPLVERPADRVDMVLYKREVPAPVAASVEVDLDQAITLANGRPGLDAA 426
QY 409 VTRDFOKAFDIANKLEVGTVHINNKTRGPDNFPFLGLKSGAGVQGIIRYSIEAMTNV 468
Db 427 VFGRDVVKIRAVRLVEGAIYIDMPRHGIGYYPFGGRKSGVFRREGIYGAVEAVTAYK 486
QY 469 STVDMK 475
Db 487 TIVFNK 493

RESULT 74
08EMH9 PRELIMINARY; PRT; 472 AA.
ID 08EMH9
AC 08EMH9:
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Succinate-semialdehyde dehydrogenase (EC 1.2.1.16).
GN OrderedLocusNames=O82863;
OS Oceanobacillus thelyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831;
RA MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkf526;
RX Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935 (2002).
CC -I- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
DR EMBL, AP004602; BAC14819.1; -.
DR HSSP, P51977; 1BXS.
DR GO, GO:0004028; F:aldehyde dehydrogenase activity; IEA.
DR GO, GO:001649; F:oxidoreductase activity; IEA.
DR GO, GO:0009013; F:succinate-semialdehyde dehydrogenase [NAD(P). .; IEA.
DR GO, GO:0008152; F:metabolism; IEA.
DR InterPro, IPR002086; Aldehyde_dehydrog.
DR InterPro, IPR010102; SSADH.
DR Pfam, PF00171; Aldehyd_1.
DR TIGRFAMs, TIGR01780; SSADH_1.
DR PROSITE, PS00070; ALDEHYD_DEHYDR_CYS; 1.
KW Complete proteome; Oxidoreductase.
KW SEQUENCE 472 AA; 51376 MW; 7DCBDB7E0F0D552 CRC64;
Query Match 29.7%; Score 711.5; DB 2; Length 472;
Best Local Similarity 34.8%; Pred. No. 1.7e-34;
Matches 167; Conservative 91; Mismatches 201; Indels 21; Gaps 7;
QY 3 KEYQNVNNGEM-KSVNIEILSPIDSSIGFVPAMTRREVDHAMKAGREALPMAALTV 61

```

0Y 8 YNNGEKSSVNOCEIISPIIDDSLSGVPMTEBVDHMMKGRZLPMALVYTERQY 67
Db 16 YINGEMITLLOEQLEVNNPATKEIFATVPGVSTKQAVDAHNEFKWUSLTUADRATK 75
0Y 68 LHRADIIERDKEIATVLAKESIKAYNASVTEVVRTADLIRYAEEGIRLTSADBGK 127
Db 76 LKMFPTLIDENKEIAMAINTKEQKPFPALEGEVYVNASPFWMYAEEBKRV-----YGE 130
0Y 128 MDASTGHKLAVIPROVGVILAPAPNPNVNLSSKIPALIGGVNWFKPTQGSGL 187
Db 131 IPAHPEPKRILVWKQEPGVMAAITPNNFPAANITRKVAPALAAAGTAAVVKASQTPYAL 190
0Y 188 VLAKEPAEAGLPACVNTITTGKSGEIGDYIYHEEVNINFTGSPVQOIGKLAG--MR 245
Db 191 KLAELHAEADIPGVINIIITGSKAMADITWMEIDGRVRSFTSGTEIGKELMASAQOMK 250
0Y 246 PYLEIGGKADAGIVLADADLDMAKQIYAGAYDSGGQSTAKRYLVYEEVADLAEKIS 307
Db 251 KVSLEIGGHAPFLVMMADLDRAVEVYVSGSKFRNAGQCCICTNRVFWOEEVYEVFEVFO 310
0Y 306 ENYAKLSVGPDPF-DNATVTPVIDDMSADPIESLVVDAROKAKELNEBFK----DGRLLT 360
Db 311 KAVGOLKATVDGFBEDGTTGVPLIDENAKSVKQGHIEDALQKGSTVLYGQKVALEBGNFIQ 370
0Y 361 P---GLFDHVTIDMKLAMEBPFPIILIRYKDAEBAVALANKSDPGLQSSVFTTRDFOKA 417
Db 371 PTVIGLANDTMLQMN---SETEGPVAPVAKFTVVDVEITERANNPFGYLAAYIFPKDISQA 427
0Y 418 FDIANKLEGVTAHNNKTRGPD--NPRILGKSGAGVQIGRISIEMTNWKSVITLDM 474
Db 428 FQISEALEVGIIGLANDGL--PSVAQAPPGGKESGIGIREGHNIEBYLEIKTISIGL 483

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2005, 10:04:24 ; Search time 167 Seconds

(without alignments)
1100.067 Million cell updates/sec

Title: US-09-868-195-12

Perfect score: 2393
Sequence: 1 LTKKEYQNYNGEMKSSVNOI.....GIRYSREANTNYKSVLDWK 475

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 80 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2382	99.5	475	5	ABP27643 Streptococcus
2	2380	99.5	475	3	AAV96743 Streptococcus
3	2373	99.2	475	5	ABP29820 Streptococcus
4	1812	75.7	474	6	ABU01538 S. pneumo
5	1812	75.7	474	6	ABP81589 Streptococcus
6	1808.5	75.6	474	8	ADK47293 Streptococcus
7	1802	75.3	474	6	ABU46050 Protein e
8	1700	71.0	475	6	ABU44392 Protein e
9	1699	71.0	439	4	AU037723 Streptococcus
10	1692	70.7	436	8	ADR95169 Novel S.
11	1641	68.6	475	5	ABP27644 Streptococcus
12	1641	68.6	475	6	ADR83939 S. pyogen
13	1639	68.5	475	6	ABU46701 Protein e
14	1397	58.4	486	6	ABU25619 Protein e
15	1388.5	58.0	486	6	ABU24815 Protein e
16	1383	57.8	482	6	ABU24149 Protein e
17	1335	55.8	472	8	ADS27503 Bacterial
18	1312	54.8	473	8	ADS27631 Bacterial
19	1244	52.0	481	8	ADS28268 Bacterial
20	1172.5	49.0	498	3	AAAG0912 Zea mays
21	1169.5	48.9	496	7	ABM74113 DNA clone
22	1168.5	48.8	496	3	AAAG32607 Arabidops
23	1168.5	48.8	496	3	AAAG9550 Arabidops
24	1168.5	48.8	496	8	ADN73355 Thale cre
25	1164.5	48.7	496	3	AAAG05755 Arabidops

26	1128.5	47.2	461	3	AAAG32608 Arabidops
27	1128.5	47.2	461	3	AAAG9551 Arabidops
28	1103.5	46.1	438	3	AAAG9552 Arabidops
29	1103.5	46.1	438	3	AAAG32609 Arabidops
30	1102.5	45.9	438	3	AAAG0913 Zea mays
31	1099.5	45.9	438	3	AAAG05754 Arabidops
32	1036.5	43.3	402	3	AAAG05755 Arabidops
33	868.5	36.3	475	6	ABU48897 Protein e
34	786.5	32.9	492	6	AAE37093 Human 913
35	784.5	32.8	293	3	AAAG40914 Human a1d
36	777	32.5	493	7	AAE39889 Human a1d
37	777	32.5	493	8	ADQ48417 Aldehyde
38	763	31.9	488	8	ADS44637 Bacterial
39	759.5	31.7	463	8	ADN18430 Bacterial
40	741.5	31.0	455	8	ADN18535 Bacterial
41	741.5	31.0	484	4	AAE79356 Coryneb
42	741.5	31.0	484	4	AAAG92676 C glutam
43	737	30.8	476	8	ADN17364 Bacterial
44	735	30.7	482	8	ADS28216 Bacterial
45	734	30.7	471	7	ADM26502 Hyperther
46	733.5	30.7	526	7	ABO78556 Pseudomon
47	728	30.4	470	8	ADS27421 Bacterial
48	726	30.3	482	4	AAU34671 E. coli c
49	726	30.3	482	6	ABU28726 Protein e
50	726	30.3	482	8	ADS45081 Bacterial
51	722	30.2	511	8	ADN46591 Bacterial
52	721.5	30.2	483	6	ABU19008 Protein e
53	721.5	30.2	488	2	AAE55135 Heat resi
54	718	30.0	488	6	ABBA7807 Listeria
55	718	30.0	488	6	ABU32914 Protein e
56	716.5	29.9	493	8	ADS29166 Bacterial
57	713.5	29.8	480	8	ADS29560 Bacterial
58	711	29.7	482	6	ABU31620 Protein e
59	711	29.7	487	7	ABO66116 Klebsiell
60	710.5	29.7	489	8	ADS29678 Bacterial
61	705.5	29.5	473	8	ADS24538 Bacterial
62	705.5	29.5	491	5	ABH08382 PUC amin
63	705.5	29.5	491	6	ABU92062 Pseudomon
64	704	29.4	480	6	ABU40285 Protein e
65	697.5	29.1	468	8	ADS28136 Bacterial
66	694	29.0	473	8	ADS29963 Bacterial
67	694	29.0	477	8	ADN26088 Bacterial
68	693	29.0	480	6	ABU41623 Protein e
69	693	29.0	480	6	ABU42152 Protein e
70	692	28.9	482	6	ABU47293 Protein e
71	691	28.9	478	8	ADS24924 Bacterial
72	690	28.8	489	6	ABU23300 Protein e
73	687	28.7	480	8	ADS22151 Bacterial
74	686	28.7	467	8	ADN25659 Bacterial
75	682.5	28.5	458	8	ADN24384 Bacterial
76	682.5	28.5	458	8	ADN21629 Bacterial
77	682.5	28.5	485	8	ADS28246 Bacterial
78	682	28.5	482	4	AAU38454 Salmonell
79	682	28.5	482	6	ABU47996 Protein e
80	680.5	28.4	460	8	ADS27648 Bacterial

ALIGNMENTS

RESULT 1
ID ABP27643 standard; protein; 475 AA.
AC ABP27643;
XX 02-JUL-2002 (first entry)
DT Streptococcus polypeptide SEQ ID NO 4462.
XX Streptococcus polypeptide SEQ ID NO 4462.
DE Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
XX Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
KW Group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

```

XX OS Streptococcus agalactiae.
XX PN WO200234771-A2.
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-GB004789.
XX PR 27-OCT-2000; 2000GB-00026333.
XX PR 24-NOV-2000; 2000GB-00028727.
XX PR 07-MAR-2001; 2001GB-00005640.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Telford J, Maignani V, Margarit Y Rost, Grandi G, Fraser C;
XX PI Tetteiln H;
XX DR WPI; 2002-352536/38.
XX DR N-PSDB; ABN68274.
XX PS
XX PT New Streptococcus protein for the treatment or prevention of infection or
XX PT disease caused by Streptococcus bacteria, such as meningitis, and for
XX PT detecting a compound that binds to the protein.
XX PS Claim 1; Page 3607; 4525pp; English.
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B
XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX CC the specification. The proteins have antibacterial and anti-inflammatory
XX CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
XX CC antibodies that bind (I) are used in the manufacture of medicaments for
XX CC the treatment or prevention of infection or disease caused by
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a
XX CC biological sample. (I) is used to determine whether a compound binds to
XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX CC used as a vaccine or diagnostic composition. The disease caused by
XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic
XX CC acid encoding (I) may be used to recombinantly produce (I) and may be
XX CC used in gene therapy. Antibodies to (I) are used for affinity
XX CC chromatography, immunoassay, and distinguishing/identifying
XX CC Streptococcus proteins
XX SQ Sequence 475 AA;

```

```

Query Match          99.5%; Score 2382; DB 5; Length 475;
Best Local Similarity 99.2%; Pred. No. 4, 5e-190;
Matches 471; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LTRKYQNYNGEMKSSVNOIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPMAALT 60
QY :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MKTEYQNYNGEMKSSVNOIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPMAALT 60
QY 61 VVERAQYLHKADIIERDKBEIATVLAKEISKAYNVSVEVTRADLIRYAAEEGRILST 120
QY :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 VVERAQYLHKADIIERDKBEIATVLAKEISKAYNVSVEVTRADLIRYAAEEGRILST 120
QY 121 SADEGGKMDASTGKHLAVIRQPVGIIVLAIPAPYNEVNLGSKTIAPALIGANVMKPEPT 180
QY :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 121 SADEGGKMDASTGKHLAVIRQPVGIIVLAIPAPYNEVNLGSKTIAPALIGANVMKPEPT 180
QY 181 QGSVSGSLVLAAPAEAGLPAGVFNITTTGSGSEIGDIYVHBEVNFNFTGSPVQGRICK 240
QY :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 181 QGSVSGSLVLAAPAEAGLPAGVFNITTTGSGSEIGDIYVHBEVNFNFTGSPVQGRICK 240
QY 241 LAGMPPIMLGKGKAGIYLAADLNLNAKQIYAGYVDYSGORCTIKKVLVVEEVADEL 300
QY :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 241 LAGMPPIMLGKGKAGIYLAADLNLNAKQIYAGYVDYSGORCTIKKVLVVEEVADEL 300
QY 301 AEKISENVAKLTVGPFDNATVTPVIDNSADFIESTLVVDARQGAKEINFEKDRGLLT 360

```

```

DB 301 AEKISENVAKLTVGPFDNATVTPVIDNSADFIESTLVVDARQGAKEINFEKDRGLLT 360
QY 361 PGLPDHNTLDMKLAEEPPFGLPIIRYKDAEBAVAINKSDPGLQSSVFRDPOKADI 420
DB 361 PGLPDHNTLDMKLAEEPPFGLPIIRYKDAEBAVAINKSDPGLQSSVFRDPOKADI 420
QY 421 ANKLEVGTVHINNKTRGPDNPPFLGLKSGAGVQIGIRYSIEAMTNVKSIVLDMK 475
DB 421 ANKLEVGTVHINNKTRGPDNPPFLGLKSGAGVQIGIRYSIEAMTNVKSIVFDMK 475
RESULT 2
AAV96743
ID AAY96743 standard; protein; 475 AA.
XX
XX AC AAY96743;
XX
XX DT 26-SEP-2000 (first entry)
XX
XX DE Streptococcus NADP-dependent glyceraldehyde-3-phosphate dehydrogenase.
XX
XX KM Group B Streptococcus; MS10; non-phosphorylating; NADP-dependent;
XX KW glyceraldehyde-3-phosphate dehydrogenase; NPGAP-3-DH; virulence;
XX KW antibiotic; vaccine; gene therapy; veterinary; mastitis.
XX
XX OS Streptococcus agalactiae.
XX
XX FH Key Location/Qualifiers
FT Misc-difference 297
FT /note="Encoded by GCW"
FT Misc-difference 439
FT /note="Encoded by CCW"
FT
PN WO200037490-A2.
XX
XX PD 29-JUN-2000.
XX
XX PF 22-DEC-1999; 99MO-GB004376.
XX
XX PR 22-DEC-1998; 98GB-00028346.
XX PR 20-JAN-1999; 99GB-00001233.
XX PR 12-APR-1999; 99GB-00001234.
XX PR 24-MAY-1999; 99GB-00008321.
XX PR 23-SEP-1999; 99GB-00022536.
XX
XX PA (MICR-) MICROSCIENCE LTD.
XX
XX PI Hughes MJG, Santangelo JD, Lane JD, Feldman R, Moore JC;
XX PI Everest P, Dobson RJ, Henwood CJ, Dougan G, Wilson RK;
XX
XX DR WPI; 2000-442636/38.
XX DR N-PSDB; AAY51293.
XX
XX PT Peptide encoded by operon including genes identified and obtained from
XX PT group B Streptococcus, for treating streptococcal infections.
XX
XX PS Example 5; Page 31-32; 32pp; English.
XX
XX CC The S. agalactiae MS14 gene encodes a protein which has homology with a
XX CC non-phosphorylating, NADP-dependent glyceraldehyde-3-phosphate
XX CC dehydrogenase (NPGAP-3-DH). Peptides derived from group B Streptococcus
XX CC (Streptococcus agalactiae) proteins encoded by genes MS4, MS10, MS11,
XX CC MS14 and/or MS16 are claimed. The peptides are useful for screening
XX CC potential drugs, or for the detection of virulence, and for the
XX CC manufacture of a medicament for use in the treatment or prevention of
XX CC infections such as focal infection (including osteomyelitis, septic
XX CC arthritis, abscesses and endophthalmitis) and urinary tract infections
XX CC caused by group B Streptococcus. The peptides and vaccines comprising the
XX CC peptides are useful in treatment of chronic mastitis, especially in cows
XX CC (i.e. for veterinary purposes)

```

Sequence 475 AA;

Query Match 99.5%; Score 2380; DB 3; Length 475;
Best Local Similarity 99.6%; Pred. No. 6.6e-190;
Matches 473; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LTKEYQNYVNGEWSKSSVNOIEILSPIDSSLGFPVPMTRBEVDHAKAGREALPAAAL 60
DB 1 LTKEYQNYVNGEWSKSSVNOIEILSPIDSSLGFPVPMTRBEVDHAKAGREALPAAAL 60
QY 61 VYERAOYLHKADIIERDKKEIATVLAKESKAYNASVTEVTRADLIRYAAEGIRLST 120
DB 61 VYERAOYLHKADIIERDKKEIATVLAKESKAYNASVTEVTRADLIRYAAEGIRLST 120
QY 121 SADEGGKMDASTGHKLAIVIRQPVGIVLAIAPYNPVNLGSKIAIPALIGGVNMFKRPPT 180
DB 121 SADEGGKMDASTGHKLAIVIRQPVGIVLAIAPYNPVNLGSKIAIPALIGGVNMFKRPPT 180
QY 181 QGSVSGVLAKAPAEAGLPAGVFNITGSGSEIGDYIVHEBEVNFNFGSTFVGQRIGK 240
DB 181 QGSVSGVLAKAPAEAGLPAGVFNITGSGSEIGDYIVHEBEVNFNFGSTFVGQRIGK 240
QY 241 LAGMRPIMLELGGKAGIVLADADLNNAAKQIVAGAYDVSQGRCTAIKRVLVVEEYADEL 300
DB 241 LAGMRPIMLELGGKAGIVLADADLNNAAKQIVAGAYDVSQGRCTAIKRVLVVEEYADEL 300
QY 301 AEKISENVAKLSVGDPEPDNATVTPVIDNSADPIESLVVDAROKAKELNEFKRDRLLT 360
DB 301 AEKISENVAKLSVGDPEPDNATVTPVIDNSADPIESLVVDAROKAKELNEFKRDRLLT 360
QY 361 PGLFDHVTLDKMLAEBEPFGLPIIRVYDAEBAVAIANKSDPGLQSSVFTTRDFQAFDI 420
DB 361 PGLFDHVTLDKMLAEBEPFGLPIIRVYDAEBAVAIANKSDPGLQSSVFTTRDFQAFDI 420
QY 421 ANKLEVGTVHINKTKGRGPDNFPPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475
DB 421 ANKLEVGTVHINKTKGRGPDNFPPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475

RESULT 3
ABP29820
ID ABP29820 standard; protein; 475 AA.
XX
AC ABP29820;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 8816.
XX
KM Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KM group A streptococcus; Streptococcus pyogenes; antibacterial;
KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
PN MO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001MO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Maignani V, Margarit y Rosi, Grandi G, Fraser C;
PI Tettein H;
XX
DR MPI; 2002-352536/38.
DR N-PSDB; ABN70451.

XX
PT New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
PS Claim 1; Page 3999; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (1), nucleic acids encoding (1), ABN6604-ABN71526 and
CC antibodies that bind (1) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (1) are used to detect Streptococcus in a
CC biological sample. (1) is used to determine whether a compound binds to
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (1) may be used to recombinantly produce (1) and may be
CC used in gene therapy. Antibodies to (1) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 475 AA;

Query Match 99.2%; Score 2373; DB 5; Length 475;
Best Local Similarity 98.9%; Pred. No. 2.5e-189;
Matches 470; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LTKEYQNYVNGEWSKSSVNOIEILSPIDSSLGFPVPMTRBEVDHAKAGREALPAAAL 60
DB 1 LTKEYQNYVNGEWSKSSVNOIEILSPIDSSLGFPVPMTRBEVDHAKAGREALPAAAL 60
QY 61 VYERAOYLHKADIIERDKKEIATVLAKESKAYNASVTEVTRADLIRYAAEGIRLST 120
DB 61 VYERAOYLHKADIIERDKKEIATVLAKESKAYNASVTEVTRADLIRYAAEGIRLST 120
QY 121 SADEGGKMDASTGHKLAIVIRQPVGIVLAIAPYNPVNLGSKIAIPALIGGVNMFKRPPT 180
DB 121 SADEGGKMDASTGHKLAIVIRQPVGIVLAIAPYNPVNLGSKIAIPALIGGVNMFKRPPT 180
QY 181 QGSVSGVLAKAPAEAGLPAGVFNITGSGSEIGDYIVHEBEVNFNFGSTFVGQRIGK 240
DB 181 QGSVSGVLAKAPAEAGLPAGVFNITGSGSEIGDYIVHEBEVNFNFGSTFVGQRIGK 240
QY 241 LAGMRPIMLELGGKAGIVLADADLNNAAKQIVAGAYDVSQGRCTAIKRVLVVEEYADEL 300
DB 241 LAGMRPIMLELGGKAGIVLADADLNNAAKQIVAGAYDVSQGRCTAIKRVLVVEEYADEL 300
QY 301 AEKISENVAKLSVGDPEPDNATVTPVIDNSADPIESLVVDAROKAKELNEFKRDRLLT 360
DB 301 AEKISENVAKLSVGDPEPDNATVTPVIDNSADPIESLVVDAROKAKELNEFKRDRLLT 360
QY 361 PGLFDHVTLDKMLAEBEPFGLPIIRVYDAEBAVAIANKSDPGLQSSVFTTRDFQAFDI 420
DB 361 PGLFDHVTLDKMLAEBEPFGLPIIRVYDAEBAVAIANKSDPGLQSSVFTTRDFQAFDI 420
QY 421 ANKLEVGTVHINKTKGRGPDNFPPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475
DB 421 ANKLEVGTVHINKTKGRGPDNFPPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475

RESULT 4
ABU01538
ID ABU01538 standard; protein; 474 AA.
XX
AC ABU01538;
XX
DT 23-OCT-2003 (revised)
DT 11-FEB-2003 (first entry)
XX

CC detecting *S. pneumoniae* in a biological sample or diagnosing *S.*
CC pneumoniae infection in a subject. The polynucleotides have antibacterial
CC activity and are useful in gene therapy

XX Sequence 474 AA;

Query Match 75.7%; Score 1812; DB 6; Length 474;
Best Local Similarity 74.1%; Pred. No. 1.8e-142;
Matches 349; Conservative 57; Mismatches 65; Indels 0; Gaps 0;

YY 5 YONVNGEKKSVNOIELISPIDSSLGFPVPMTEEVDMKAGREAPAMALTYER 64
DB 4 YONLVNGKMKSSBOEITTYSPINOELGTVPAMTQTEADAMQAAAPALPAMALSAVER 63
YY 65 AOVTHKADIIERDKERIATVLAKEISKAVNASVTEVTRADLIRAAEGIRLSTADE 124
DB 64 AAVLHTAALIERDKKEIGITIAKEVAKGKKAIGEVTRADLIRYAAEBGIRITQOME 123
YY 125 GGRMDASTGHKLAVIRROPVGIATAPVNPVNLSSKIAAPALIGANNVMPKPTQGSV 184
DB 124 GGGFEATSKNKLAVRRBPVGIATAPVNPVNLSSKIAAPALIGANNVMPKPTQGSV 183
YY 185 SGLVAKAFAPAEAGLPAGVFNITIGRSGEIGDYIVHEBVNFINTGSTPVGQRIGLAGM 244
DB 184 SGLLAKAFEEAGCPAGVFNITIGRSGEIGDYIIEHKVNFINTGSTPIGERIGLAGM 243
YY 245 RPIMLELGGKAGIATADADLNAKQIVAGAYDYGQRCTAIKRVLVVEVADLAEKI 304
DB 244 RPIMLELGGKDALVEDADLEHAQKQIVAGAFSYSGQRCTAIKRVLVESVADKLATLL 303
YY 305 SENVAKLVSVDPEPDNATVPVIDNSADFIESLVVAROKGAKELNEFRDGRLLRGLP 364
DB 304 QEBVSKLTVGDPDPDNADITPVIDNASADFIWGLIEDAQEKQAALTPIKREGNLLMPVLP 363
YY 365 DHTYLMKLAWEPPFPIPIIRVKDAEVAIAANKSDFGLOSSVFTDPFOKAFDIANKL 424
DB 364 DQYTKMKAKAEEPPFPIPIIRVAVSEBAIAFANESSEFLOSSVFTDPFOKAFDIANKL 423
YY 425 EVGTVINNKTRGPDNPFPLGLKSGAGVQIGRYSIEAMTNVKSIVLDMK 475
DB 424 EVGTVINNKTRGPDNPFPLGLKSGAGVQIGRYSIEAMTNVKSIVFDVK 474

RESULT 6

ADK47293 standard; protein; 474 AA.

XX ADK47293;

XX 20-MAY-2004 (first entry)

XX Streptococcus pneumoniae protein, Seq ID No 3808.

XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.

XX Streptococcus pneumoniae.

XX US6699703-B1.

XX 02-MAR-2004.

XX 26-MAY-2000; 2000US-00583110.

XX 02-JUL-1997; 97US-0051553P.

XX 12-MAY-1998; 98US-0085131P.

XX 30-JUN-1998; 98US-00107433.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;

XX WPI; 2004-212399/20.

XX N-PDB; ADK44632.

XX New nucleic acid molecules and polypeptides useful for diagnosing,
PT preventing and treating pathological conditions resulting from bacterial
PT infection, e.g. Streptococcus pneumoniae infection, and in drug
PT screening.

XX Disclosure; SEQ ID NO 3808; 301pp; English.

CC The invention relates to isolated Streptococcus pneumoniae nucleic acids
CC and polypeptides. The nucleic acids and proteins are useful for
CC diagnosing, preventing and treating pathological conditions resulting
CC from bacterial infection, such as *S. pneumoniae* infection. These may also
CC be used for drug screening procedures. The present sequence represents a
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
CC data for this patent did not appear in the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.

XX Sequence 474 AA;

Query Match 75.6%; Score 1808.5; DB 8; Length 474;
Best Local Similarity 73.7%; Pred. No. 3.6e-142;
Matches 350; Conservative 59; Mismatches 65; Indels 1; Gaps 1;

YY 1 LTRKYONVNGEKKSVNOIELISPIDSSLGFPVPMTEEVDMKAGREAPAMALTYER 60
DB 1 LTRKYONLVNGKMKSSBOEITTYSPINOELGTVPAMTQTEADAMQAAAPALPAMALSAVER 59
YY 61 YVERAQLYKADIIERDKERIATVLAKEISKAVNASVTEVTRADLIRYAAEGIRLST 120
DB 60 ATERAAYLHKTALIERDKKEIGITIAKEVAKGKKAIGEVTRADLIRYAAEBGIRITG 119
YY 121 SADBEGKMDASTGHKLAVIRROPVGIATAPVNPVNLSSKIAAPALIGANNVMPKPT 180
DB 120 QMEGGFPAASKNKLAVRRBPVGIATAPVNPVNLSSKIAAPALIGANNVMPKPT 179
YY 181 QSSVSGVLAFAEAGLPAGVFNITIGRSGEIGDYIVHEBVNFINTGSTPVGQRIGK 240
DB 180 QSSISGLLAKAFEEAGCPAGVFNITIGRSGEIGDYIIEHKVNFINTGSTPIGERIGR 239
YY 241 LAGMPRIMLELGGKAGIATADADLNAKQIVAGAYDYGQRCTAIKRVLVVEVADL 300
DB 240 LAGMPRIMLELGGKDALVEDADLEHAQKQIVAGAFSYSGQRCTAIKRVLVESVADKL 299
YY 301 AEKISENVAKLSVGPDPDNATVPVIDNSADFIESLVVAROKGAKELNEFRDGRLLT 360
DB 300 ATTLQEBVSKLTVGDPDPDNADITPVIDNASADFIWGLIEDAQEKQAALTPIKREGNLLM 359
YY 361 PGLFPHVTLDMKLAWEPPFPIPIIRVKDAEVAIAANKSDFGLOSSVFTDPFOKAFDI 420
DB 360 PVLFPQVTKDMKLAWEPPFPIPIIRVAVSEBAIAFANESSEFLOSSVFTDPFOKAFDI 419
YY 421 ANKLEVGTVHINKTRGPDNPFPLGLKSGAGVQIGRYSIEAMTNVKSIVLDMK 475
DB 420 AEKLEVGTVHINKTRGPDNPFPLGLKSGAGVQIGRYSIEAMTNVKSIVFDVK 474

RESULT 7

ABU46050 standard; protein; 474 AA.

XX ABU46050;

XX 19-JUN-2003 (first entry)

XX Protein encoded by prokaryotic essential gene #31577.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Streptococcus pneumoniae.

XX MO200277183-A2.

CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation (9)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from Wipo at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 475 AA;

Query Match 71.0%; Score 1700; DB 6; Length 475;
Best Local Similarity 66.9%; Pred. No. 4, 2e-133;
Matches 318; Conservative 75; Mismatches 82; Indels 0; Gaps 0;

QY 1 LTRKQYNYNGEWSVNOIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPAAWALT 60
DB 1 MKQYQYNYNGEWSVNOIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPAAWALT 60
QY 61 YVERAQLHAKADIRDEBEATVLAKEISKAYNASVTEVTRADLIRYABEGRLST 120
DB 61 YVERAQLHAKADIRDEBEATVLAKEISKAYNASVTEVTRADLIRYABEGRLST 120
QY 121 SAEDEGKMDASTGHLAVIRROPVGIIVLAIPANYVNVNLSGSLIAGNVWFKPPT 180
DB 121 EVDGGSFPAASKKIAVVRREBVGVLAIISPNTVNVNLGSLIAGNVWFKPPT 180
QY 181 QGSVSGVLAKAPAEAGLPAGVNTITGSGEIGDYIVHEEYVNFINTGSPVVGORIK 240
DB 181 QGSISGLLAAEAPAGLPAGVNTITGSGEIGDYIVHEEYVNFINTGSPVVGORIK 240
QY 241 LAGMPRIMELGSKADGIVLADADLNAKQIVAGYVDSGRCIAIKRYLVVEVADBL 300
DB 241 MAGMRIMELGSKASIVLEDDDELTAKNITAGAFYSGQRCIAIKRYLVVEVADBL 300
QY 301 AEKISNVAKLSVGDPEFNATVTPVTDNSADPTESLVVDARQKAKEINEFRDGRILT 360
DB 301 VEKIREKVALITIGNEPDADITPLIDTYSADIVBGLINDANDKGAALFETIKREGNLIC 360
QY 361 PGHFDHTLDMKLAWEPEFPIPIIRVDAEBAVIAKNSDGLQSSVFTEDFOKAPDI 420
DB 361 PLEFDKVTTLDMKLAWEPEFPIPIIRVDAEBAVIAKNSDGLQSSVFTEDFOKAPDI 420
QY 421 ANKLEVGTVHINNTKRGPDNPFPLGLKSGAGVQIRYSIEAMTVKSIIVLDMK 475
DB 421 AEOLVEGVTHINNTKRGPDNPFPLGLKSGAGVQIRYSIEAMTVKSIIVLDMK 475

RESULT 9

AAU37723 standard; protein; 439 AA.

AAU37723;

14-FEB-2002 (first entry)

Streptococcus pneumoniae cellular proliferation protein #152.

Anti-sense; prokaryotic cellular proliferation protein; antibiotic;
antibacterial; drug design.

OS Streptococcus pneumoniae.

XX WO200170955-A2.

XX 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2001; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GD;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX N-PSDB; AAS55582.PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 13316; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from Wipo at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 439 AA;

Query Match 71.0%; Score 1699; DB 4; Length 439;
Best Local Similarity 74.5%; Pred. No. 4, 5e-133;
Matches 327; Conservative 55; Mismatches 57; Indels 0; Gaps 0;

QY 37 MTRBEVDHAKAGREALPAAWALTVERAQLHAKADIRDEBEATVLAKEISKAYNA 96
DB 1 MTRBEVDHAKAGREALPAAWALTVERAQLHAKADIRDEBEATVLAKEISKAYNA 96
QY 97 SVEVVRTRADLIRYABEGRLSTSADEGKMDASTGHLAVIRROPVGIIVLAIPANY 156
DB 97 SVEVVRTRADLIRYABEGRLSTSADEGKMDASTGHLAVIRROPVGIIVLAIPANY 156
QY 157 VNLSSKTIAPALIGNVVFKPPTGSGVGLVLAFAEAGLPAGVNTITGSGEIGDY 216
DB 157 VNLSSKTIAPALIGNVVFKPPTGSGVGLVLAFAEAGLPAGVNTITGSGEIGDY 216
QY 217 IVEHEEVNFINTGSPVVGORIKLAGMPRIMELGSKADGIVLADADLNAKQIVAG 276
DB 217 IVEHEEVNFINTGSPVVGORIKLAGMPRIMELGSKADGIVLADADLNAKQIVAG 276
QY 277 YDYGORCTAIRVAVVEVADLAKEISNVAKLSVGDPEFNATVTPVTDNSADPTE 336
DB 277 YDYGORCTAIRVAVVEVADLAKEISNVAKLSVGDPEFNATVTPVTDNSADPTE 336
QY 336 YDYGORCTAIRVAVVEVADLAKEISNVAKLSVGDPEFNATVTPVTDNSADPTE 336
DB 336 YDYGORCTAIRVAVVEVADLAKEISNVAKLSVGDPEFNATVTPVTDNSADPTE 336

Qy	337	LIVDAPRGAKXINFEKRDGRLLTQGLFPHWLTDLMAEHPFGPILPIIRVYDAEAEVA	396
Db	301	LIEDQKXKAOQLTIKKEGNLTUMPLFPOVTKDKMVAHEPPGVPVLPIIRVASVEEATA	360
Qy	397	IANKSDPGLQSSVFTRDPFOKAPDIANKLEVGVTHINNKTKGRGDPNPFGLKSGAGVQG	456
Db	361	FANESEFGLQSSVFTRNDFFKKAPEIAKELEVGVTHINNKTKGRGDPNPFGLKSGAGVQG	420
Qy	457	IRYSTEAMTNVKSITVLDMK	475
Db	421	IKYSIEAMTNVKSIVEDVK	439
RESULT 10			
ID	ADR95169	standard; protein; 436 AA.	
XX	ADR95169;		
AC			
DT	16-DEC-2004	(first entry)	
XX			
De	Novel S. pneumoniae protein sequence, SEQ ID 3804.		
XX			
KW	Meningitis; bacteraemia; pneumonia; otitis media; vaccine;		
KW	bacterial infection.		
XX			
OS	Streptococcus pneumoniae.		
XX			
EN	US6800744-B1.		
XX			
PD	05-OCT-2004.		
XX			
PF	30-JUN-1998;	98US-001077433.	
XX			
XX	02-JUN-1997;	97US-0051553P.	
PR	12-MAY-1998;	98US-0085131P.	
XX			
PA	(GENO-) GENOME THERAPEUTICS CORP.		
XX			
PI	Doucette-Stamm LA, Bush D;		
XX			
DR	WPI, 2004-697205/68.		
DR	N-PSDB; ADR92566.		
XX			
PT	New isolated nucleic acid encoding a Streptococcus pneumoniae		
PT	polypeptide, useful for diagnosing, preventing and/or treating		
PT	pathological conditions resulting from the bacterial infection.		
XX			
PS	Disclosure; SEQ ID NO 3804; 151pp; English.		
XX			
CC	The invention relates to an isolated nucleic acid comprising a sequence		
CC	encoding a Streptococcus pneumoniae ADR9136polypeptide, or its		
CC	fragments, with any of 9 fully defined sequences (appearing as ADR94308,		
CC	ADR944800, ADR944837, ADR94969, ADR95253, ADR95642, ADR95682,		
CC	ADR96079) or any of the fully defined sequences appearing as ADR91705,		
CC	ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR9366, ADR9650 or		
CC	ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide		
CC	sequences, or at least 40, 60 or 300 consecutive nucleotides, which is		
CC	hybridisable under high stringency conditions to the nucleotide sequence.		
CC	The nucleic acids and proteins are chosen from 5206 disclosed sequences.		
CC	Also included are a recombinant expression vector comprising the isolated		
CC	nucleic acid cited above operably linked to a transcription regulatory		
CC	element, a cell comprising the recombinant expression vector and a probe		
CC	comprising at least 20 consecutive nucleotides of the nucleotide		
CC	sequences as cited above. The methods and compositions of the present		
CC	invention are useful for the diagnosis, prevention and/or treatment of		
CC	pathological conditions resulting from bacterial infection by		
CC	Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and		
CC	otitis media. The present sequence is one of the 2603 disclosed S.		
CC	pneumoniae protein sequences. Note: The sequence data for this patent did		
CC	not form part of the printed specification, but was obtained in		
CC	electronic format directly from USPTO at		

CC	seqdata.uspfco.gov/sequence.html?PDocID=6800744B1
XX	Sequence 436 AA;
XX	Query Match 70.7%; Score 1692; DB 8; Length 436;
XX	Best Local Similarity 74.9%; Pred. No. 1.7e-132;
XX	Matches 326; Conservative 53; Mismatches 56; Indels 0; Gaps 0
QY	41 EVDHAKAGREALPAWALTYYERAOYLHKADIIERDKSEIATVLAKESISKAYNASVTE 1000
DB	2 EADEGQAAARALPAARALSAIERAAYLHKTALIERDKKEIGTILAKETAKGKAIGE 61
QY	101 VVATADLIYYAAEEGRILSTADSEGRKMDASTGHKLAVIRQPVGIYLAIPVNYPNLS 160
DB	62 VVATADLIYYAAEEGRILITGQAMEGGGFEEAASKNDLAVVREPVGIYLAIPVNYPNLS 121
QY	161 GSKTIAPALIGSNVWMEKPPQTQSGVGLYAKAFAPAGLVFTITGRGSEIGDYIYEH 220
DB	122 ASKTIAPALIGSNVWMEKPPQTQSGISGLLAKAFEBAGIPACVFTITGRGSEIGDYIIEH 181
QY	221 EEVNFINFTGSPVPGORIGKLAKMFPIMLEIGSGDAGIVLADADLDNNAKOIVAGAYDYS 280
DB	182 KEVNFIFNFTGSPPIGERRIGRLAKMFPIMLEIGSGDAAVLVEDADLEHAAKOIVAGAFSYS 241
QY	281 GORCTIKKIVLVVEBEVADLAKIESBNAXKISVGDPPNATVTFVIDINDNSADPLESIVVD 340
DB	242 GORCTIKKIVLVVESVADLAKITLQEEBSKILVGDPPFNATITVIDINDNSADPLTWGLIED 301
QY	341 AROKGAKEINFEKRDGRLLTPGLPDHVTLDMLKAMEBPPGJILPIIRVKDAEBAVAIANK 400
DB	302 AOEKEKQALTPIKRBENLMPVLFPQVTKDKKVAKEBPPGVLDIIRVASVEEALAPANE 361
QY	401 SDFGLQSVFTNDFOKAPDIANKLEVGTVHINNTKGRGPNFPPLGLKSGAGVQGIKYS 460
DB	362 SERGLQSVFTNDFOKAFETIAKLEVGTVHINNTKQGRGPNFPPLGVKSGAGVQGIKYS 421
QY	461 IEAMTVKSIIVLDMK 475
DB	422 IEAMTVKSIIVEDVK 436
XX	RESULT 11
XX	ABP27644
XX	ID ABP27644 standard; protein; 475 AA.
XX	AC ABP27644;
XX	DT 02-JUL-2002 (first entry)
XX	DE Streptococcus polypeptide SBQ ID NO 4464.
XX	KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX	KM group A streptococcus; Streptococcus pyogenes; antibacterial;
XX	KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX	OS Streptococcus pyogenes.
XX	PN WO200234771-A2.
XX	PD 02-MAY-2002.
XX	PF 29-OCT-2001; 2001WO-GB004789.
XX	PR 27-OCT-2000; 2000GB-00026333.
XX	PR 24-NOV-2000; 2000GB-00028727.
XX	XX 07-MAR-2001; 2001GB-00005640.
XX	PA (CHIR-) CHIRON SPA.
XX	PA (GENO-) INST GENOMIC RES.
XX	PI Telford J, Maignani V, Margarit Y Rosl, Grandi G, Frazer C;
XX	Tetteijn H;

DR MPI: 2002-352536/38.
DR N-PSDB; ABN68275.
PT New Streptococcus protein for the treatment or prevention of infection or
XX disease caused by Streptococcus bacteria, such as meningitis, and for
XX detecting a compound that binds to the protein.
PS Claim 1; Page 3607; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and anti-inflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
XX Sequence 475 AA;
SQ
Query Match 68.6%; Score 1641; DB 5; Length 475;
Best Local Similarity 65.7%; Pred. No. 3.5e-128;
Matches 312; Conservative 76; Mismatches 87; Indels 0; Gaps 0;
1 LTKXYONYNGEMKSSVNOIEILSPIDBSLGFVPMTRREYVDHAKGREALPMAALT 60
1 LAKQYKNLVNGEMKSSVNOIEILSPIDBSLGFVPMTRREYVDHAKGREALPMAALT 60
61 YVERAOYLHKADIIERDKSEIATVLAKEISKAAYNASVTEVTRADLIRYAEGRIST 120
61 YVERAOYLHKADIIERDKSEIATVLAKEISKAAYNASVTEVTRADLIRYAEGRIST 120
61 YVERAOYLHKADIIERDKSEIATVLAKEISKAAYNASVTEVTRADLIRYAEGRIST 120
121 SAEDEGKMDASGHKLAIVRRBPVGLVLAIPYNNVNLGSKTAPALTAGVNVALKPPT 180
121 EYLBEGSGFPAASKKKAIVRRBPVGLVLAIPYNNVNLGSKTAPALTAGVNVALKPPT 180
181 QGSVSGLVYAKAFABEAGIPAGVFNITGRGSEIGDIYVEHEEVNFINTGSTPVQGRICK 240
181 QGSISGLLLAFAFAEAGIPAGVFNITGRGSEIGDIYVEHEEVNFINTGSTPVQGRICK 240
241 LAGMRPIMELGKKGAGIYVADADLADNAKQIVAGAYDYSQGRCTAIRKVLVEYADEL 300
241 LAGMRPIMELGKKGAGIYVADADLADNAKQIVAGAYDYSQGRCTAIRKVLVEYADEL 300
301 AKKISENVAKLSVGPDPNATVTPVIDNSADPISLVDAROKGAKELNEFKDGRLLT 360
301 AAEIKTLVKLSVGMPEDDADITPILDTSAADFVEGLIKDADKGTALTATNREGNLIS 360
361 PGLFHDVTLDMKLAWEPEPGLPIPIRYKDAEAAVIAKNSDFGLQSSVFTDFOKAFDI 420
361 PGLFHDVTLDMKLAWEPEPGLPIPIRYKDAEAAVIAKNSDFGLQSSVFTDFOKAFDI 420
421 ANKLEVGTVHINKTGRGPDNPFPLGLKSGAGVQGRIRYSIEAMTVKSIIVLDMK 475
421 ANKLEVGTVHINKTGRGPDNPFPLGLKSGAGVQGRIRYSIEAMTVKSIIVLDMK 475
421 ABLQLEVGTVHINKTGRGPDNPFPLGLKSGAGVQGRIRYSIEAMTVKSIIVLDMK 475
RESULT 12
ADR83939
ID ADR83939 standard; protein; 475 AA.
XX
XX ADR83939;
AC
XX
DT 02-DEC-2004 (first entry)

XX
XX S. pyogenes hyperimmune system reactive antigen SpY1371.
DE hyperimmune serum reactive antigen; vaccine; anticaline.
KW Streptococcus pyogenes.
OS WO2004078907-A2.
XX
XX 16-SEP-2004.
XX
XX 02-MAR-2004; 2004WO-BP002087.
XX
XX 04-MAR-2003; 2003EP-00450061.
XX
XX (INTE-) INTERCELL AG.
XX
XX Meinke A, Nagy E, Winkler B, Gelbmann D;
PI MPI: 2004-653698/63.
DR N-PSDB; ADR83789.
XX
XX New isolated nucleic acid molecules encoding hyperimmune serum-reactive
PT antigens from Streptococcus pyogenes, useful for diagnosing, preventing
PT and treating S. pyogenes infections.
XX
XX Claim 11; SEQ ID NO 207; 145pp; English.
XX
XX This invention describes a novel nucleic acid molecule encoding a
CC hyperimmune serum reactive antigen or its fragment from Streptococcus
CC pyogenes. The nucleic acid molecule or hyperimmune serum-reactive antigen
CC or its fragment are useful for the manufacture of a pharmaceutical
CC preparation, especially for the manufacture of a pharmaceutical
CC addition, the hyperimmune serum reactive antigen or fragment is used for
CC the isolation and/or purification and/or identification of an interaction
CC partner of the hyperimmune serum reactive antigen or its fragment, for
CC the generation of a peptide (e.g. anticalins) binding to the antigen or
CC fragment, or for the manufacture of a functional nucleic acid selected
CC from aptamers and spligenomers. The nucleic acid molecule may also be used
CC for the manufacture of functional ribonucleic acids, such as ribozymes,
CC antisense nucleic acids and siRNA. ADR83733-ADR84189 represent S.
CC pyogenes hyperimmune serum reactive antigens, fragments and the encoding
CC polynucleotide described in the invention.
XX
XX Sequence 475 AA;
SQ
Query Match 68.6%; Score 1641; DB 8; Length 475;
Best Local Similarity 65.7%; Pred. No. 3.5e-128;
Matches 312; Conservative 76; Mismatches 87; Indels 0; Gaps 0;
1 LTKXYONYNGEMKSSVNOIEILSPIDBSLGFVPMTRREYVDHAKGREALPMAALT 60
1 LAKQYKNLVNGEMKSSVNOIEILSPIDBSLGFVPMTRREYVDHAKGREALPMAALT 60
61 YVERAOYLHKADIIERDKSEIATVLAKEISKAAYNASVTEVTRADLIRYAEGRIST 120
61 YVERAOYLHKADIIERDKSEIATVLAKEISKAAYNASVTEVTRADLIRYAEGRIST 120
61 YVERAOYLHKADIIERDKSEIATVLAKEISKAAYNASVTEVTRADLIRYAEGRIST 120
121 SAEDEGKMDASGHKLAIVRRBPVGLVLAIPYNNVNLGSKTAPALTAGVNVALKPPT 180
121 EYLBEGSGFPAASKKKAIVRRBPVGLVLAIPYNNVNLGSKTAPALTAGVNVALKPPT 180
181 QGSVSGLVYAKAFABEAGIPAGVFNITGRGSEIGDIYVEHEEVNFINTGSTPVQGRICK 240
181 QGSISGLLLAFAFAEAGIPAGVFNITGRGSEIGDIYVEHEEVNFINTGSTPVQGRICK 240
241 LAGMRPIMELGKKGAGIYVADADLADNAKQIVAGAYDYSQGRCTAIRKVLVEYADEL 300
241 LAGMRPIMELGKKGAGIYVADADLADNAKQIVAGAYDYSQGRCTAIRKVLVEYADEL 300
301 AKKISENVAKLSVGPDPNATVTPVIDNSADPISLVDAROKGAKELNEFKDGRLLT 360
301 AAEIKTLVKLSVGMPEDDADITPILDTSAADFVEGLIKDADKGTALTATNREGNLIS 360

Qy 361 PGLFDVHTLDMKLAKEEPPEGLPIIRIVYDAEEAAVANKSPGLOSSVFPRDQKFDI 420
Db 361 PVLFDAHTTDMRLAKEEPFGVLPPIIRVTTVEALKINSEBYGLQASIPFTNPKKAFGI 420
Qy 421 ANKLDEYCTHAINKTGRGPDNPFFGLKSSGAGVCGIRISTEAMTNVKSITVLDMK 475
Db 421 ABOLEGVSTHANKTGRGTDNPFGLAKKSGAGVGVKRSLEAMTVKSVVFDIQ 475

CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the largest prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC [ftp.wipo.int/pub/published_pct_sequences](http://www.wipo.int/pub/published_pct_sequences)
XX
SQ Sequence 475 Aa;


```

DB 251 GMLPWLLEIGKDAIVLEDAINLETTAKSIVSAGVYSQRCAYRVLYMKVADLELVE 310
QY 303 KISENVAKISVGDPPDNATVTPVIDNSADFTESLVVDARQKAKELNEFKRDGRLITPG 362
DB 311 LVTKKREKELKVGPPFDVITPLIDRKADVYQTLIDDDAIKCATLIVNKKREINLMPY 370
QY 363 LFDHVTLMKLAWEPPGPIIPRIYKDAEENVALINKSGDFGUSVFRDPOKAFDIAN 422
DB 371 LFDNVTADMKIAWEEPPGPVLPRIYKSDAELELNRSBYGLQSAVFTEENHDAFYIAN 430
QY 423 KLEVGTVHINNKTGRGPDNFPFLGLKSGAGVQIGIRYSIEAMTNVKSIVLDM 474
DB 431 KLDVGTVOYNNKREGRGPDHFPFLGTSSGMDGQIGIRYSIEAMTRHKSIVLNL 482

RESULT 17
ADS27503
ID ADS27503 standard; protein; 472 AA.
XX
AC ADS27503;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #16536.
XX
KM Recombinant DNA construct; transformed plant; improved plant property;
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KM pathogen tolerance; pest tolerance; plant disease resistance;
KM cell cycle pathway modification; plant growth regulator;
KM homologous recombination; seed oil yield; protein yield; carbohydrate;
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KM bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 16536; 122pp; English.
XX
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX CC The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests.

```

```

CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 472 AA;
XX
Query Match 55.8%; Score 1335; DB 8; Length 472;
Best Local Similarity 55.7%; Pred. No. 1.3e-102;
Matches 263; Conservative 80; Mismatches 127; Indels 2; Gaps 1;

QY 5 YGNVNGEMK--SSVNOJELISPPIDSSLGFPVPMTRREVDHAKAGREALPAMALITVY 62
DB 1 YKPYLNGEMRESSSETTEIIPSPYHBYIGVQVATIRGEVDRAIASAKAQSMBASIQ 60
QY 63 ERAQVILHKAADIIERDKSEIATVLAKEISKAYNASVTEVYRTADLIRYAEBSGIRLSTSA 122
DB 61 DRAKVLVYKADDELVMQDEIADIIMKEVGKGYKDAKKEVYRTADLIRYIEALMHGHS 120
QY 123 DEGGKMDASTGKLAIVIRROPVGIYLAIPYVPYVNLSSGKIAPLIGNVVMEKPPYOG 182
DB 121 NMGDSFPFGGTGSKLAIIRAPGVVLAIPFVYPVNLISAKLAIPALIMGNNAVIFRPATOG 180
QY 183 SVSGVLAKAFPAEALPGVFNTITRGSEIGDYVEHEVGFNIPFGSTPVGQIIGKLA 242
DB 181 AISGIRKWEALHKAALPGVLVNAVATGKSVIGDYVEHEGIMMVSTFGTNTGKHLAKKA 240
QY 243 GMRPIMLEIGKDAIGVLAADLDNAAKOIVAGAVDYSGORCTAIKRVLVVEEVADELAE 302
DB 241 SMIPVLVEIGKDPRIAREDDADLOPAANHIVSGAISYSGORCTAIKRVLVHENVADLEVD 300
QY 303 KISENVAKISVGDPPDNATVTPVIDNSADFTESLVVDARQKAKELNEFKRDGRLITPG 362
DB 301 ILKQVAVELSVGSPQDSTIVPLIDDKSADFQGLVDDAVEKGATVIGNKKERNLIYPT 360
QY 363 LFDHVTLMKLAWEPPGPIIPRIYKDAEENVALINKSDPELGSSVFRDPOKAFDIAN 422
DB 361 LIDHVTLEMKIAWEEPPGPVLPRIYKSDAELELNRSBYGLQSAVFTEENHDAFYIAN 420
QY 423 KLEVGTVHINNKTGRGPDNFPFLGLKSGAGVQIGIRYSIEAMTNVKSIVLDM 474
DB 421 KIETGSGVQINGRTERGPDHFPPIYGVKSGMGAGQIRKSLSESTREKIVVNLN 472

RESULT 18
ADS27631
ID ADS27631 standard; protein; 473 AA.
XX
AC ADS27631;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #16664.
XX
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX

```

PD 18-DEC-2003.
 XX 20-FEB-2003; 2003US-00369493.
 PF
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 PS
 PS Claim 1; SEQ ID NO 16664; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX
 XX
 SQ Sequence 473 AA;

Query Match 54.8%; Score 1312; DB 8; Length 473;
 Best Local Similarity 55.2%; Pred. No. 1.1e-100;
 Matches 261; Conservative 77; Mismatches 133; Indels 2; Gaps 1;

QY 5 YONYNGEMK--SSVNOIEILSPIDSSIGFVPAMTREVVDHAKMGREALPMALATVY 62
 DB 1 YNPFVVGEMRESSSGETIISSEYLYHEVIGQVQITCGEVDKAIASAKEXMAWAKA 60
 QY 63 ERAQYVHKRADIIERKEIATVLAKEISKAVASTVEVTRTDLRVAAEEIRLSTQA 122
 DB 61 DRAKTVYKWADELVMNMODEIADIMEVGVGADAKKEVVRVDFRYYTEALHMHGSS 120
 QY 123 DEGGKMDASTGHLAVIRROPVGIIVLAIPYNYVNLSSKIPALIGGVVWFKEPTOG 182
 DB 121 NMGDSFPGTSGTSLAITIQRAPLGVALIAPFNYPVNLAKAPALIMENAVIFKATOG 180
 QY 183 SVSGVLVAFAPABAGPAGVFNITTRGSGEIGDYIVHEBEVNFINTGSTPVGQRIKLA 242
 DB 181 AISGIMVBAIHKAGPGKGLVWVATGRGSAIGDYLVGHGIMNVSTFGNTGNTKHLAKTA 240
 QY 243 GMRPIMLEIGKDGAGIVLADADIDNNAKQIVAGAYVSGRCATIRVVLVEVADBLAE 302
 DB 241 AMIPIVLEIGKDPGIVREBDADLDQADANHIVSGAFSFGORCTAIRVIVHENVADBLVD 300

QY 303 KISENVAKLSVGDPEPNATVTPVIDNSADPESLIVDARQKAKELNEFRKRGDLTPG 362
 DB 301 LKQAVAEISLSPKESKOSTIVRLIDDKSADPFQGLVDDAVEKGAATIVIGNRERLIVPT 360
 QY 363 LFDHYTLDMKLAMEEPFGPIIPIIRVKDAEVAIAKNSDGLQSVTFTRPQAFDIAN 422
 DB 361 LIDHYKEMKVAWEPFGPIIPIIRVSSDEQAIERIANSEBFGLOASVETKIKINAKFAIAN 420
 QY 423 KLEVGTVHNNKTVGSGPDPNPPFLGKSGAGVQVGRYSIEAMTVNKSIVLDMK 475
 DB 421 KIETGVOINGRTKTERGDPHFPFPIGVKSGMGAGQIRKSLBSMTREKTVVLNFK 473

RESULT 19
 ADS28268
 ID ADS28268 standard; protein; 481 AA.
 XX
 AC ADS28268;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #17301.

XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

OS Bacteria.

PN US2003233675-A1.

PD 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 17301; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or

AC ABM74113;
XX
DT 17-OCT-2003. (first entry)
XX
DE DNA clone originating in barley containing SNP sequence #523.
XX
KM Barley; single nucleotide polymorphism, SNP; genotype-phenotype analysis.
OS Hordeum vulgare.
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
XX
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYN1-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284bp; Japanese.
XX
XX The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences
XX
SQ Sequence 496 AA;

Query Match 48.9%; Score 1169.5; DB 7; Length 496;
Best Local Similarity 49.8%; Pred. No. 9.1e-89;
Matches 238; Conservative 84; Mismatches 143; Indels 13; Gaps 5;

QY 5 YONYVNGEMKSSVN--QIELSPIDDSIGFVPMTRREVDHMKRGREALPMALTYV 62
DB 16 YKYVADGEMRPASISGTVTAIVNPTROTQYRVQACTQEEVNKYWDAAKVAQKAWATPIM 75
QY 63 ERAQYLKRAADIIERDKKEIATLAKESKAYNASYTEVVRATDLIRYAAEGIRLSTSA 122
DB 76 KGAELHKAALILKEKTPILAECLVKEINKPADAVSEVVRSGDLYSYTAEQVRLGE- 134
QY 123 DEGGKM--DASTG--HKLAVIRQPGVIALAIYNTYPVNLSGSKIAPALIGNVNMF 176
DB 135 ---GKLVSDSPFGERNRNYCCLSSKVPILGVLAIPFPNYPVNLAIVSKIGPALTAGNSLYI 191
QY 177 KPTQGSVGLVYAKAPAEAGLPAVENVITTGSGSEIGDVIYHEEVNFINFTGSPVQO 236
DB 192 KPTQGAVALAHVHCFHLNAGFPKGLISCTGKSGSEIGDFTMHDPVNCISFTGG-DTGI 250
QY 237 RIGKLGAEMPIIMLEGGKDGAGIVLADADLDNAKOIVAGAYDYSGORCTAIRKVLVEEV 296
DB 251 AISKAGWVPLQWELGKDACTVLEDDADLDVLAANIVKGFSTSGRCRAVKKVLLMEV 310

QY 297 ADELAKEISENVAKLVGDPDPDNATVTPYIDNSADFIESTLVVDAROKGAKELNEPKRG 356
DB 311 ADAVEKVNAAKAKLVGPPEDDCITTPVTESSANFIEGLVMDAKKCATTCQETRRRG 370
QY 357 RLITRGLEFDHVTLLDMKLMWEEPPGPIPLPIRVKDAEBAVALNKSDFGLQSSVFTTRDFOK 416
DB 371 NLIWPLLDHVPDMRIAMEEPPGFVLPVIRINSVVEEGIHHCNANSGFLQCGCVFTRDINK 430
QY 417 AFDIANKELEVGTVAHNNKTGRGPDNPPLGLKSGAGVQGIYSIEAMTNVKSIVLDM 474
DB 431 AIMISDAMETGTVOINSAPARGPDHFPFQGLKDSGISQGITNSINMTKVKSVTINTL 488

RESULT 22
AAG32607
ID AAG32607 standard; protein; 496 AA.
XX
AC AAG32607;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 39371.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-01223180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130691P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 14-MAY-1999; 99US-0134768P.
PR 18-MAY-1999; 99US-0134941P.
PR 19-MAY-1999; 99US-0135124P.
PR 20-MAY-1999; 99US-0135153P.
PR 21-MAY-1999; 99US-0135529P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0143277P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.

PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155148P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157353P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 12-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 48.8%; Score 1168.5; DB 3; Length 496;
Best Local Similarity 50.3%; Pred. No. 1.1e-88;
Matches 240; Conservative 84; Mismatches 142; Indels 11; Gaps 5;

PR	22-JUL-1999	99US-01450855
PR	22-JUL-1999	99US-01450872
PR	22-JUL-1999	99US-01450899
PR	22-JUL-1999	99US-01451929
PR	23-JUL-1999	99US-01451452
PR	23-JUL-1999	99US-01452188
PR	23-JUL-1999	99US-01455244
PR	26-JUL-1999	99US-01452766
PR	27-JUL-1999	99US-01459138
PR	27-JUL-1999	99US-01455183
PR	28-JUL-1999	99US-01455919
PR	02-AUG-1999	99US-01465386
PR	02-AUG-1999	99US-01463888
PR	02-AUG-1999	99US-01465899
PR	03-AUG-1999	99US-01470388
PR	04-AUG-1999	99US-01472040
PR	04-AUG-1999	99US-01473022
PR	05-AUG-1999	99US-01471932
PR	05-AUG-1999	99US-01477808
PR	06-AUG-1999	99US-01473030
PR	06-AUG-1999	99US-01474162
PR	09-AUG-1999	99US-01478399
PR	09-AUG-1999	99US-01479359
PR	10-AUG-1999	99US-01481712
PR	11-AUG-1999	99US-01483192
PR	12-AUG-1999	99US-01483414
PR	13-AUG-1999	99US-01485652
PR	15-AUG-1999	99US-01486849
PR	17-AUG-1999	99US-01491758
PR	18-AUG-1999	99US-01493426
PR	20-AUG-1999	99US-01497223
PR	20-AUG-1999	99US-01497233
PR	20-AUG-1999	99US-01499922
PR	23-AUG-1999	99US-01499022
PR	23-AUG-1999	99US-01499302
PR	25-AUG-1999	99US-01505662
PR	26-AUG-1999	99US-01508849
PR	27-AUG-1999	99US-01510652
PR	27-AUG-1999	99US-01510662
PR	27-AUG-1999	99US-01510802
PR	30-AUG-1999	99US-01513030
PR	31-AUG-1999	99US-01514382
PR	01-SEP-1999	99US-01519302
PR	07-SEP-1999	99US-01522632
PR	10-SEP-1999	99US-01530702
PR	13-SEP-1999	99US-01537582
PR	15-SEP-1999	99US-01540182
PR	16-SEP-1999	99US-01540392
PR	22-SEP-1999	99US-01547792
PR	22-SEP-1999	99US-01551332
PR	23-SEP-1999	99US-01554862
PR	24-SEP-1999	99US-01556592
PR	28-SEP-1999	99US-01564582
PR	29-SEP-1999	99US-01565662
PR	04-OCT-1999	99US-01571132
PR	05-OCT-1999	99US-01577532
PR	06-OCT-1999	99US-01576652
PR	07-OCT-1999	99US-01580292
PR	08-OCT-1999	99US-01582322
PR	12-OCT-1999	99US-01583632
PR	13-OCT-1999	99US-01592232
PR	13-OCT-1999	99US-01592842
PR	13-OCT-1999	99US-01592852
PR	14-OCT-1999	99US-01593392
PR	14-OCT-1999	99US-01593312
PR	14-OCT-1999	99US-01596372
PR	14-OCT-1999	99US-01596382
PR	18-OCT-1999	99US-01598642
PR	21-OCT-1999	99US-01607412
PR	21-OCT-1999	99US-01607672

PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161932P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0161942P.

Query Match	48.8%;	Score 1168.5;	DB 3;	Length 496;
Best Local Similarity	50.3%;	Pred. No. 1.1e-88;		
Matches 240;	Conservative 84;	Mismatches 142;	Indels 11;	Gaps 5

QY	5	QUNYNGWK--SSNYQJELISPIDDSSLGVNPATREVDHAKAGBALPAMALTY	62
Db	16	YKYADGEMKTSSSSKSVAILMNPATRKQYRYQACTQEBVAVMELASQKSMKTPIM	75
QY	63	BRQYLIHAKAADIIEEDKEEIAVTLAKEISKVYNASVTAVR7ADLIIRYAAEGILSTSA	122
Db	76	KRAELIHQAAILKDNKAPMAESLVKEIAKPAKOSVTEVNSGDLISYCAEBGVAL--L	132
QY	123	DEGKM--DASTGH--KLAVIRROPVGIVALAIAPYNPVNISGSKIAPALIGANVMFK	177
Db	133	GEKFLSDSPFGNDRKCYCLTSKIPLGVLVLAIPFPNPVNLAVSKIAIPALTAGNSLVK	192
QY	178	PPTQSSVGLVLAKAFAPAGPAGVFNITIGRGSTIGDIYIHEHEVNPINFTSGTPVQR	237
Db	193	PPTQAVSCLMHVHCFHLGFPFKGISCTIGSGSEIGDFTMHPPVNCISFTGGS-DTGIS	251
QY	238	IGTLAKGKPMLELGGKAGIVLADADLDNNAKOIVAAAYDVSQGRCTAIRKVLVEEVA	297
Db	252	ISKRAKGMPLQMELOGKACIVLDDADLDVLAASNIIKGFSTSGRCITAKVYLVME5YA	311
QY	298	DELAKEISENVAKISVGDPPFNATVTPVIDNSADFIESLVDAROKGAKELNEFKRDR	357
Db	312	DELVEKVAKAKVAKLTVGPPENSDITAVVSSSANFIEGLVMDAKEKATFCQEKYRBN	371
QY	358	LITPLFPHVYLTDMKLAEEPPGPILPIIRKYADEAVALANKSDPGIQSSVFTDPQKA	417
Db	372	LIMPILLDNVPPDMKIAEPPGPVAVPLARINSVEEINHCHCASNFGIGQCFYTDINKA	431
QY	418	FDIANKLEVTGVHINKTKRGDPNPPFLGLKSGAGVQIGIRYSIAMTNVKSIVLDM	474
Db	432	ILISDAMETGYQINSAPARGPDHPFPGGLKDSIGSGGVNINSILMTKVKTTVINL	488
RESULT 24			
ID	ADN73355	standard; protein; 496 AA.	
XX	ADN73355;		
XX	15-JUL-2004	(first entry)	
DE	Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 1250.		
XX			
KW	plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;		
KW	animal feed product; thale cress; cell wall biosynthesis;		
XX	nitrogen metabolism; carbon metabolism.		
OS	Arabidopsis thaliana.		
XX			
PN	WO2004035798-A2.		
XX			

Db	76	KTELHHRQAAILKONKAPMAESLVEIKKPAKDSYTEVVRSDGLSYCABEBSVRI--L	132
Oy	123	DEGGKM--DASTGH--KLAVIRROPGVIVLAIPANYPNVLGSKIAIAPLIGANVMFK	177
Db	133	GEQKFLSDSPFGNDRTKYCLTNSKPLGVVLAIPPNYPVNLAVSKIAIPALLAGNSLVLK	192
Oy	178	PPTQGSVSGLVAKAPAEAGLPAQVENVITTGSGSEIGDYIVEHEEVPNFINTGSTPEGOR	237
Db	193	PPTQGAVSGTLAHVHCFHLAGFRKGLISCTTGKSGEIGDPLTTHPANCISFTGQ-DTGIS	251
Oy	238	IGKLAGEPIMELGSGKQAGIVLADADLDMAAKOIVAGAYDSGQRTAIKRVLVVEEVA	297
Db	252	ISKKGMIPLQWELGSGKQACIVLDDADLVLVSNIIKGGSPYSGQCTAVKQVVLVWESVA	311
Oy	298	DELAEKISENVAKLSVGDPEFNATVTVTVIJDNSADFIESLVDVARQKAGEKELNEFRDGE	357
Db	312	DELVEKVVAKVAKLITVGPPEENSDTLAVVSESSANPIEGLVMDAKKGAATFCGEYKREGN	371
Oy	368	LITPGLFDHVTILDMKLAMEEPFGPILPIIRVXDAEFAVALIANKSDPGLQSSVFTRPQKA	417
Db	372	LITPILLDNVBRDPMRIAMEEPFGPVVPVPIRINSVEGINHCNANSFGLQGCVEFKDINXA	431
Oy	418	FDIANKLEVGTHINNKTRGPDNPNPFLGKSGAGVQGIRISIEAMTNVKSIVLDM	474
Db	432	ILISAMEGTVOINSAPARGDHPFQGLKDSGQSVTNSINIMTKVKTIVINDL	488

RESULT	26
AAG32608	
ID	AAG32608 standard; protein; 461 AA.
XX	
AC	AAG32608;
XX	
DT	17-OCT-2000 (first entry)
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 39372.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KV	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128334P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	28-APR-1999; 99US-0130891P.
PR	30-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	07-MAY-1999; 99US-0132487P.
PR	11-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.

PR	14-MAY-1999	9905-0134418P
PR	14-MAY-1999	9905-0134411P
PR	14-MAY-1999	9905-01344221P
PR	14-MAY-1999	9905-0134370P
PR	18-MAY-1999	9905-0134376P
PR	19-MAY-1999	9905-0134491P
PR	20-MAY-1999	9905-0135124P
PR	21-MAY-1999	9905-0135553P
PR	24-MAY-1999	9905-0135621P
PR	25-MAY-1999	9905-0136021P
PR	27-MAY-1999	9905-0136392P
PR	28-MAY-1999	9905-0136782P
PR	01-JUN-1999	9905-0137222P
PR	03-JUN-1999	9905-0137528P
PR	04-JUN-1999	9905-0137502P
PR	07-JUN-1999	9905-0137724P
PR	08-JUN-1999	9905-0138069P
PR	10-JUN-1999	9905-0138540P
PR	14-JUN-1999	9905-0138647P
PR	14-JUN-1999	9905-0139119P
PR	16-JUN-1999	9905-0139452P
PR	16-JUN-1999	9905-0139453P
PR	17-JUN-1999	9905-0139492P
PR	18-JUN-1999	9905-0139454P
PR	18-JUN-1999	9905-0139456P
PR	18-JUN-1999	9905-0139457P
PR	18-JUN-1999	9905-0139458P
PR	18-JUN-1999	9905-0139459P
PR	18-JUN-1999	9905-0139460P
PR	18-JUN-1999	9905-0139461P
PR	18-JUN-1999	9905-0139462P
PR	18-JUN-1999	9905-0139463P
PR	23-JUN-1999	9905-0140034P
PR	24-JUN-1999	9905-0140695P
PR	28-JUN-1999	9905-0140823P
PR	29-JUN-1999	9905-0140991P
PR	30-JUN-1999	9905-0141887P
PR	01-JUL-1999	9905-0141842P
PR	01-JUL-1999	9905-0142154P
PR	02-JUL-1999	9905-0142059P
PR	06-JUL-1999	9905-0142390P
PR	08-JUL-1999	9905-0142803P
PR	09-JUL-1999	9905-0143220P
PR	12-JUL-1999	9905-0142977P
PR	13-JUL-1999	9905-0143542P
PR	14-JUL-1999	9905-0143524P
PR	15-JUL-1999	9905-0144005P
PR	16-JUL-1999	9905-0144085P
PR	16-JUL-1999	9905-0144086P
PR	19-JUL-1999	9905-0144325P
PR	19-JUL-1999	9905-0144331P
PR	19-JUL-1999	9905-0144332P
PR	19-JUL-1999	9905-0144333P
PR	19-JUL-1999	9905-0144343P
PR	19-JUL-1999	9905-0144335P
PR	20-JUL-1999	9905-0144352P
PR	20-JUL-1999	9905-0144632P
PR	20-JUL-1999	9905-0144684P
PR	21-JUL-1999	9905-0144641P
PR	21-JUL-1999	9905-0144681P
PR	21-JUL-1999	9905-0145088P
PR	22-JUL-1999	9905-0145087P
PR	22-JUL-1999	9905-0145089P
PR	22-JUL-1999	9905-0145192P
PR	23-JUL-1999	9905-0145187P
PR	23-JUL-1999	9905-0145218P

PR	18-JUN-1999	9905-01394559
PR	18-JUN-1999	9905-01394610
PR	18-JUN-1999	9905-01394619
PR	18-JUN-1999	9905-01394622
PR	18-JUN-1999	9905-01394633
PR	18-JUN-1999	9905-01394750
PR	18-JUN-1999	9905-01394763
PR	21-JUN-1999	9905-01396117
PR	22-JUN-1999	9905-01396999
PR	23-JUN-1999	9905-01403539
PR	23-JUN-1999	9905-01403544
PR	24-JUN-1999	9905-01406599
PR	28-JUN-1999	9905-01408823
PR	30-JUN-1999	9905-01409891
PR	30-JUN-1999	9905-01412878
PR	01-JUL-1999	9905-01415422
PR	01-JUL-1999	9905-01421815
PR	02-JUL-1999	9905-01420255
PR	06-JUL-1999	9905-01425390
PR	09-JUL-1999	9905-01428033
PR	09-JUL-1999	9905-01428202
PR	12-JUL-1999	9905-01428777
PR	14-JUL-1999	9905-01433422
PR	14-JUL-1999	9905-01433624
PR	15-JUL-1999	9905-01440059
PR	16-JUL-1999	9905-01440859
PR	16-JUL-1999	9905-01440866
PR	19-JUL-1999	9905-01443252
PR	19-JUL-1999	9905-01443311
PR	19-JUL-1999	9905-01443322
PR	19-JUL-1999	9905-01443333
PR	19-JUL-1999	9905-01443434
PR	19-JUL-1999	9905-01443435
PR	19-JUL-1999	9905-01443522
PR	20-JUL-1999	9905-01443832
PR	20-JUL-1999	9905-01443843
PR	21-JUL-1999	9905-01443814
PR	21-JUL-1999	9905-01450862
PR	22-JUL-1999	9905-01450889
PR	22-JUL-1999	9905-01450887
PR	22-JUL-1999	9905-01450892
PR	23-JUL-1999	9905-01451522
PR	23-JUL-1999	9905-01452182

PR	23-JUL-1999	99US-01455224P
PR	26-JUL-1999	99US-0145276P
PR	27-JUL-1999	99US-0145513P
PR	27-JUL-1999	99US-0145518P
PR	27-JUL-1999	99US-0145519P
PR	28-JUL-1999	99US-0145591P
PR	28-JUL-1999	99US-0145591P
PR	02-AUG-1999	99US-0146386P
PR	02-AUG-1999	99US-0146388P
PR	02-AUG-1999	99US-0146389P
PR	03-AUG-1999	99US-0147038P
PR	04-AUG-1999	99US-0147204P
PR	04-AUG-1999	99US-0147302P
PR	05-AUG-1999	99US-0147192P
PR	05-AUG-1999	99US-0147260P
PR	06-AUG-1999	99US-0147303P
PR	06-AUG-1999	99US-0147416P
PR	09-AUG-1999	99US-0147439P
PR	09-AUG-1999	99US-0147935P
PR	10-AUG-1999	99US-0148171P
PR	11-AUG-1999	99US-0148319P
PR	12-AUG-1999	99US-0148341P
PR	13-AUG-1999	99US-0148561P
PR	13-AUG-1999	99US-0148684P
PR	16-AUG-1999	99US-0149368P
PR	17-AUG-1999	99US-0149175P
PR	18-AUG-1999	99US-0149426P
PR	18-AUG-1999	99US-0149722P
PR	20-AUG-1999	99US-0149723P
PR	20-AUG-1999	99US-0149929P
PR	23-AUG-1999	99US-0149930P
PR	23-AUG-1999	99US-0149930P
PR	25-AUG-1999	99US-0150566P
PR	26-AUG-1999	99US-0150884P
PR	27-AUG-1999	99US-0151065P
PR	27-AUG-1999	99US-0151066P
PR	27-AUG-1999	99US-0151080P
PR	30-AUG-1999	99US-0151303P
PR	31-AUG-1999	99US-0151438P
PR	01-SEP-1999	99US-0151930P
PR	07-SEP-1999	99US-0152363P
PR	10-SEP-1999	99US-0152073P
PR	13-SEP-1999	99US-0153758P
PR	15-SEP-1999	99US-0154018P
PR	16-SEP-1999	99US-0154039P
PR	22-SEP-1999	99US-0154779P
PR	22-SEP-1999	99US-0155139P
PR	24-SEP-1999	99US-0155486P
PR	24-SEP-1999	99US-0155659P
PR	28-SEP-1999	99US-0156458P
PR	29-SEP-1999	99US-0156596P
PR	04-OCT-1999	99US-0157117P
PR	05-OCT-1999	99US-0157753P
PR	06-OCT-1999	99US-0157665P
PR	07-OCT-1999	99US-0158029P
PR	08-OCT-1999	99US-0158232P
PR	12-OCT-1999	99US-0158369P
PR	13-OCT-1999	99US-0159293P
PR	13-OCT-1999	99US-0159294P
PR	14-OCT-1999	99US-0159568P
PR	18-OCT-1999	99US-0159584P
PR	21-OCT-1999	99US-0160761P
PR	21-OCT-1999	99US-0160767P
PR	21-OCT-1999	99US-0160770P
PR	21-OCT-1999	99US-0160814P
PR	21-OCT-1999	99US-0160815P
PR	22-OCT-1999	99US-0160980P
PR	22-OCT-1999	99US-0160981P

PR	25-OCT-1999	99US-0161089P
PR	25-OCT-1999	99US-0161044P
PR	25-OCT-1999	99US-0161405P
PR	25-OCT-1999	99US-0161405P
PR	25-OCT-1999	99US-0161359P
PR	26-OCT-1999	99US-0161359P
PR	26-OCT-1999	99US-0161360P
PR	26-OCT-1999	99US-0161361P
PR	28-OCT-1999	99US-0161920P
PR	28-OCT-1999	99US-0161992P
PR	28-OCT-1999	99US-0161933P
PR	29-OCT-1999	99US-0162142P

Query Match	47.2%;	Score 1128.5;	DB 3;	Length 461;
Best Local Similarity	51.6%;	Pred. No. 2.2e-85;		
Matches 230;	Conservative 78;	Mismatches 129;	Indels 9;	Gaps 4

QY	3	VPMATREEDVDMKMGKREGLPAMALTVYEEBAQVYLHKRAADIIEEDKXEIATVTLAKEISKA	93
Db	12	VQCTSEEVNANWELKASKHOKSMATTPJMKRBEHLHKRAALILKONKAPMAESLYKEIAKP	71
QY	94	YNASVTEVVRTADLIRYAABEGIRLSTSADEGGKM--DASTGH--FLAVIRROPVGIYL	148
Db	72	AKQSVTEVVRSGDLISYCAEEGVRI--LGEQKFLSDSPFGNDRTYXCILTSKIPLGVLL	128
QY	149	AIAPNYTPVNLSSSKTAPRLIGGNVYMKRPPTQSGVSGVLAKFPALAGLPAGFNTITG	208
Db	129	AIPPEFVYPVNLAVSKIAAPLLIAGNSLVLPKPTQAGVSGCLMHVHCFHLAGEPKGLISCTIG	188
QY	209	RGEISDGYIYEHEEVNFIFNTGSTPEVGORIGKLAGMREIMLELGKDAGIVLADADLDA	268
Db	189	KGEIIGDFLTMHPANVNCISFTGG-DTGISISKKAGMPLQMEELGKQACIYLDADIDLIV	247
QY	269	AKQIVAGAYDVGORTAIKRVLVVEEYADBLAEKISENVAKLSVGPDPNATVTPYIDD	328
Db	248	ASNIIRIKGSEYSGQRTAVKAVVLIWESVADDELVEKVKAKVAKLTVGPEEENSIDITAVVSE	307
QY	329	NSADPLESHVVDAROKGAKELNEFPKRDRLTLPGLPFDHVTLDMLAMEBPPRGILPIIRV	388
Db	308	SSANPFIISGLVMDAKKEGATFCQEYIKRBGNLLPFLDLIDVNRDPMKRIAMEBPPRGVPPVLR	367
QY	389	KDAEEVAALANKSDPGLQSGSVFTRPFOKAFDIANKLEGVTHINNKTGGRGPDNPFGLGK	448
Db	368	NSVBEGBNMQMANSFPGLOSCVFTKQINKALLISDMAMGTVQINSAPARGDHPFPFGLK	427
QY	449	GSGAGVQVIGIRYSTEAMTNVKSIVLDM	474
Db	428	DSGIGSGQVYTNISINLMTKVKYTVINL	453
RESULT 27			
AAG49551			
ID	AAG49551 standard; protein; 461 AA.		
XX	AAG49551;		
XX	18-OCT-2000 (first entry)		
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 62698.		
XX	protein identification; signal transduction pathway; metabolic pathway;		
KM	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
XX	Arabidopsis thaliana.		
OS	EP1033405-A2.		
PN	06-SEP-2000.		
PD	25-FEB-2000; 2000EP-00301439.		
PF	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		

PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140033P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140659P.
PR 29-JUN-1999; 99US-0140623P.
PR 30-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154799P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.

PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159295P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 47.2%; Score 1128.5; DB 3; Length 461;
Best Local Similarity 51.6%; Pred. No. 2,2e-85;
Matches 230; Conservative 78; Mismatches 129; Indels 9; Gaps 4;

QY 34 VPAMTREVVDHAKMGREALPMAALTYVRAQYILKAADIIEKDEIATVLAKEISKA 93
DB 12 VQACTOEYVNAVMEILAKSAQKSWAKTPLMKRAELHKAALIKDNKAPMAESLVEKEIAKP 71
QY 94 YNASTVEVVRTADLIRYAEEGIRLSTADEGKM--DASTGH--KLAIVIRQPGIVL 148
DB 72 AKDSVTEVRSGLDLSYCAEEGVRI--LGEKFFLSDSPFGNDRTKYCLTSKIPLGVL 128
QY 149 AIAPIVYVNLGSGKIAIPALIGNVVMFKPPTQSGVSGVLAKAFAPAGLPAQVNTITG 208
DB 123 AIPFPNYPNVLAVSKIAPIALINGNSLVLPPTQGANVSCILMHVCHFLAAGFPKGLISCTIG 188
QY 209 RGSRTIGDYIVHEHEVNFINTGSTPVGORIGKLGNRPIMLEIGCKDAGIVLADADLDNA 268
DB 189 KSEIIGDFLTMHPAVNVCISFTGG-PTGISISKAGMIFLOMELGCKDAGIVLADADLDLV 247
QY 269 ANQIYAGADYSGQRTAIKRLVLYVEVVADELAEKISENVAKLSVCDPPDNATVTPVID 328
DB 248 ASNIIKGGSYSGORTAVKVVLYVMSVDELVEKAKVAKLTGVPRENSDITAVASE 307
QY 329 NSADFTESLVVDAROKAKELNEFKRDLTPGLPDHYTLMLKLAWEPPGPILPIIRV 388
DB 308 SSANITEGLVMDAKKEGATFCOYRKREGNLIMPLDLDNVRPDMRLAMEBPPGIVFVLXI 367
QY 389 KDAEEVAVALNKSDPGLQSGSVFTPDFOKAFDIANKLEVGTVHINNKTGRGPNFPPLIGLK 448
DB 368 NSEVEGINHCNASNFGLOCGCVFTKDIINKAILLSDAMEITQVONSPAPAGPDHPFPGIGLK 427
QY 449 GSGAGVQGIIRYSEIEMTNTKSYVLDM 474
DB 428 DSGISGQGVTSINIMTAKVKTIVINL 453

RESULT 28
AAG49552
ID AAG49552 standard; protein; 438 AA.

XX XX
AC AAG49552;
XX 18-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SEQ ID NO: 62699.
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62699.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0133256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134321P.
PR 14-MAY-1999; 99US-0134321P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.

[illegible]

Oy		261	GORCAIRKRVLYVEEVADPLAKRISNNVAKLSVGDFPDNATVPYIDDNSADPTESLIVD	340
Db		237	GQRCAVRAKVYLWVESVADLEVERKVAQAKLTGVPEEENSDDTAAYVSESSANPIEGILVD	296
Oy		341	AROKAKEINERKRDGRLLTPGLFDHVLDMKLAMEEPFGPFIPIIRVKDAEEAVAIANK	400
Db		297	AKKKATCCQEYFKREGNLWPILLDNVRPDMRIAMEEPPGFPVYPVARINSVEGINHCNA	356
Oy		401	SDFGIQSSVFTRDFQKAFDIANKLEVGTVINNKTRGGDPNPFLGLKSGAGVGGIRYS	460
Db		357	SNFGLQGCVFTKDINKAILIISDAMERTVOINSAPARGEDHFPGIKDSGISGGVTNS	416
Oy		461	IEMTNVKSIVLDM	474
Db		417	IMMTKVKTYINV.L	430
<hr/>				
RESULT_29				
AAG32609				
ID		AAG32609	standard; protein; 438 AA.	
XX		AAG32609;		
AC		17-OCT-2000	(first entry)	
XX				
DT				
XX			Arabidopsis thaliana protein fragment SEQ ID NO: 39373.	
XX				
KM			Protein identification; signal transduction pathway; metabolic pathway;	
KW			hybridisation assay; genetic mapping; gene expression control; promoter;	
XV			termination sequence.	
XX				
OS			Arabidopsis thaliana.	
XX				
EN			EP1033405-A2.	
XX				
PD			06-SEP-2000.	
PF			25-FEB-2000; 2000EP-00301439.	
XX				
PR		25-FEB-1999;	99US-0121825P.	
PR		05-MAR-1999;	99US-0123180P.	
PR		09-MAR-1999;	99US-0123548P.	
PR		23-MAR-1999;	99US-0125788P.	
PR		25-MAR-1999;	99US-0126264P.	
PR		29-MAR-1999;	99US-0126785P.	
PR		01-APR-1999;	99US-0127462P.	
PR		06-APR-1999;	99US-0128234P.	
PR		08-APR-1999;	99US-0128714P.	
PR		16-APR-1999;	99US-0129845P.	
PR		19-APR-1999;	99US-0130077P.	
PR		21-APR-1999;	99US-0130449P.	
PR		23-APR-1999;	99US-0130510P.	
PR		23-APR-1999;	99US-0130891P.	
PR		28-APR-1999;	99US-0131449P.	
PR		30-APR-1999;	99US-0132048P.	
PR		04-MAY-1999;	99US-0132407P.	
PR		04-MAY-1999;	99US-0132484P.	
PR		05-MAY-1999;	99US-0132485P.	
PR		06-MAY-1999;	99US-0132486P.	
PR		06-MAY-1999;	99US-0132487P.	
PR		07-MAY-1999;	99US-0132863P.	
PR		11-MAY-1999;	99US-0134256P.	
PR		14-MAY-1999;	99US-0134218P.	
PR		14-MAY-1999;	99US-0134219P.	
PR		14-MAY-1999;	99US-0134221P.	
PR		14-MAY-1999;	99US-0134370P.	
PR		18-MAY-1999;	99US-0134376P.	
PR		19-MAY-1999;	99US-0134941P.	
PR		20-MAY-1999;	99US-0135124P.	
PR		21-MAY-1999;	99US-0135353P.	
PR		24-MAY-1999;	99US-0135629P.	
PR		25-MAY-1999;	99US-0136021P.	

PR	7-MAY-1999;	99US-01363832P
PR	26-MAY-1999;	99US-01372222P
PR	01-JUN-1999;	99US-01372222P
PR	03-JUN-1999;	99US-01375282P
PR	04-JUN-1999;	99US-01375020P
PR	07-JUN-1999;	99US-01377247P
PR	08-JUN-1999;	99US-01380094P
PR	10-JUN-1999;	99US-01386400P
PR	10-JUN-1999;	99US-01388477P
PR	14-JUN-1999;	99US-01391119P
PR	16-JUN-1999;	99US-01394532P
PR	16-JUN-1999;	99US-01394543P
PR	17-JUN-1999;	99US-01394542P
PR	18-JUN-1999;	99US-01394545P
PR	18-JUN-1999;	99US-01394546P
PR	18-JUN-1999;	99US-01394610P
PR	18-JUN-1999;	99US-01394588P
PR	18-JUN-1999;	99US-01394592P
PR	21-JUN-1999;	99US-01398177P
PR	22-JUN-1999;	99US-01398033P
PR	22-JUN-1999;	99US-01400533P
PR	23-JUN-1999;	99US-01400534P
PR	24-JUN-1999;	99US-01400655P
PR	28-JUN-1999;	99US-01400823P
PR	29-JUN-1999;	99US-01400951P
PR	30-JUN-1999;	99US-01412671P
PR	01-JUL-1999;	99US-01418424P
PR	01-JUL-1999;	99US-01421544P
PR	02-JUL-1999;	99US-01420555P
PR	06-JUL-1999;	99US-01423300P
PR	08-JUL-1999;	99US-01428030P
PR	09-JUL-1999;	99US-01429270P
PR	12-JUL-1999;	99US-01429777P
PR	19-JUL-1999;	99US-01443335P
PR	19-JUL-1999;	99US-01443331P
PR	19-JUL-1999;	99US-01443332P
PR	19-JUL-1999;	99US-01443333P
PR	19-JUL-1999;	99US-01443334P
PR	19-JUL-1999;	99US-01443335P
PR	19-JUL-1999;	99US-01443336P
PR	21-JUL-1999;	99US-01445086P
PR	21-JUL-1999;	99US-01445088P
PR	22-JUL-1999;	99US-01445087P
PR	22-JUL-1999;	99US-01445089P
PR	22-JUL-1999;	99US-01445090P
PR	22-JUL-1999;	99US-01445132P
PR	23-JUL-1999;	99US-01445145P
PR	23-JUL-1999;	99US-01445218P
PR	23-JUL-1999;	99US-01445224P
PR	26-JUL-1999;	99US-01452176P
PR	27-JUL-1999;	99US-01452133P
PR	27-JUL-1999;	99US-01452181P
PR	27-JUL-1999;	99US-01452198P
PR	28-JUL-1999;	99US-01452951P
PR	02-AUG-1999;	99US-01463386P
PR	02-AUG-1999;	99US-01463388P
PR	02-AUG-1999;	99US-01463389P
PR	03-AUG-1999;	99US-01467038P

PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142877P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148314P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148656P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-015659P.
PR 29-SEP-1999; 99US-0156458P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158022P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160800P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.1%; Score 1102.5; DB 3; Length 438;
Best Local Similarity 50.6%; Pred. No. 3e-83;
Matches 220; Conservative 78; Mismatches 126; Indels 11; Gaps 4;

QY 46 MKAGREALPAMALTYERAQYVHKAADIIERDEKEIATVLAKEISKANNAVTEVVRKA 105
1 MPAKVAQAKRAMATPLMKRAELHKAALIKHEKRAELAECLVLEIAKPAKDAVSEVVRSG 60
DB 106 DLIRVAEEGIRISTNADDEGKK---DASTG---HKLAVRRQPVGIVLAIAFYNPVNL 159
61 DLVSYAEAGVRLIGE---GKLLVSDSFPGNERNKYCLSSKIPLVVLAIPFPNPVNL 116
QY 160 SSGKIAPALIGVWVWFKPTQGSVSGVLVAKAFAEAGLPAGVNTTGRSGRIQYIYE 219
117 AUSKIPALIAAGNALVUKPTQAVAAALHWHGFLHAGFPKGLISCVTKGSGIDPLTM 176
DB 220 HEEVNFINTGSTPVGQRIGKLAGMRPIMLELGKQDAGIVLADADLDNNAKOIVAGAYDY 279
177 HPGVNCISFTGG-DTGIAISKKAGWVPLQWELGKQACILEADADDVSNIVKGFSGY 235
QY 280 SSGRCIAIKRVLVVEVADELAKEISENNAKLSVGPFPNATVPTVYIDNSADPISLIV 339
236 SSGRCIAVAVVLMESIDAVVQKVNAKAKLKVGPREDSDITPVVTSSANFISGLVM 295
QY 340 DAROKAGALENEFKRGRLITPGLPFDHYTLDMKLAMEBPPGPIITIRVYKAEAAVALAN 399
296 DAKKGAATTCQETRRRGNTLWPLLDVHRPDMRIAMEBPPGVLPIRINSVEEGIHHCN 355
DB 400 KSDPGLQSSVFTDFQAFDIANKLEVGTVHINNKGKRGPDNFPFLGLKSGAGVQIRY 459
356 ASNFGIQQGCIFFTRDIKAILISAMETGYQINSARAGDHRPFGGLDSDGIGSGGGIN 415
QY 460 STEAMTNVKSIVIDM 474
DB 416 SINMTKVKXSTVINL 430

RESULT 31

AAG05754 standard; protein; 438 AA.

AAG05754;

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 2277.

KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX XX
OS Arabidopsis thaliana.
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.

Db 357 SNFGLOGCVFTKDINKAIIISDAMEGTCTQINSAPARGDPHFQGLKSGISGCVTMS 416
Qy 461 IEAMTNVKSIVLDM 474
Db 417 INLMTKVKTTVINL 430

RESULT 32
AAG05755
ID AAG05755 standard; protein; 402 AA.
XX AAG05755;
AC AAG05755;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 2278.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131499P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147193P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.

PR	12-AUG-1999;	99US-0146341P.
PR	13-AUG-1999;	99US-0146565P.
PR	13-AUG-1999;	99US-0146684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.

Qy	137	AVIRROPGVILAIAPYANYPVNLSSKTAAPLITGNVYMFPPPOGVSGLVIAKAPAEA	156
	:	:	:
Db	58	CLTSLKPIELGVLAIPFNPVLAIVASKAPALIAJNSLVLEKPPQGVASCLEHMHCFHIA	117
	:	:	:
Qy	197	GLPAGVNTITGRGSEIGDVIYEHEBNVNFETSTPVGORIGLIAQMRPIMLEIGKDA	256
	:	:	:
Db	118	GFPRGLISCTIGKSEIGDPLTMHPAVNCISFTGG-DGISISKKAAMIPLOMELGGKDA	176
	:	:	:
Qy	257	GIYLAADIDMNAKOIVGAYDYSGORCTAIVRYLVVEBVDELAEKISENVAKLISVGD	316
	:	:	:
Db	177	CYLDHDDDLVANSNITIKGFSYSQRCIAKVAVLAHVESVDELVEYKAKVAALIVGP	236
	:	:	:
Qy	317	FDNATTPYIDNSADPIESLVVDAROKAGAELEFKRDGELLTPGLFDHVTIDMKLAME	376
	:	:	:
Db	237	EENSIDITAVSSESNAPFIEGLVMDAKEGATFCQBYKREGULIWPILLIDVNRPMRIAME	296
	:	:	:
Qy	377	EPRGPELPIIRVYDMEEAVALANKSDPGLQSSVFPDRQKAFDIAKLELVGTYHINNTG	436
	:	:	:
Db	297	EPGPVAVPVARINSVEGINHCNASNFGIQCVPFKIDINKKILISDMETGTVOINSAPA	356
	:	:	:
Qy	437	RGPDNFPFLGLKSGAGVGQIVRSIEMANTYKSVILDM	474
	:	:	:
Db	357	RGPDPHPFQGLKSGIGSGVINSINLMTKXTVYINL	394
	:	:	:

RESULT 33
 ABU4897
 ID ABU4897 standard; protein; 475 AA.
 AC
 ABU4897;
 XX
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #34424.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Ureaplasma urealyticum.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 XX
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Hasselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-029926/02.
 DR N-PSDB; ACA52767.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 76821; 1766bp; English.
 XX
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 475 AA:

Query Match 36.3%; Score 868.5; DB 6; Length 475;
Best Local Similarity 39.7%; Pred. No. 1.2e-63;
Matches 190; Conservative 95; Mismatches 177; Indels 17; Gaps 5;

QY 5 YQNVYNGEMKSNVNOJELISPIDSSLGFPVPMATRE-EVDHAKKAGREALPMAALTYVE 63
DB 3 YKLINGAFADAEKKEKPVNPSNNQIAYVPHNHEINTEIPENHIALKCKDPIPKX 62
QY 64 RAQYLKKAADIIRDEKEIATVLAKEISKAYNSVTEVVTADLIRYAEGIRLSTSA- 122
DB 63 KCLLLKADKDLDEHKOELAOITSTELAKGKOSLFEVSSADYLETVEYQKLMQKP 122
QY 123 --DEGGMDASTGH---KLAVIRQPVGIVLAIAPYNYPVNLGSKIAPALIGNVNMF 176
DB 123 IFDE-----TVHNVKVAATYTRIPVGIVLAICPNYVINLLISLAPALVSGNSLTV 175
QY 177 KPTQGSVSGVLAKAPAEAGLPAGVFNITTTGSGSELDYIVHEEYVNFNFGSTPVQ 236
DB 176 KPSTQSLTIGIRSELVHEVGFPGVNNCTITTEARITGDLVNNKVKALSTFGGPRVGN 235
QY 237 RICKLAGMRPIMLELGGKAGIYLAADADLNAKQIYAGAYDSGORCTAIKRVLVEEV 296
DB 236 HIMEITSKISLVELEGKDPALVDDDFELANNEIVKGYGSGQRCTAIKRVFVSHKN 295
QY 297 ADRLAEKISENVAKLSVGPDPDNATVTVPIDNSADFIESLVDARQGA--KELNEPKK 354
DB 296 HDLLVNLINKKVALTVGLPQNPITPLINSNSLKNLSLVEDAIKKGAIYHOKIVNNE 355
QY 355 DGRLLTPGLFDVHTLDMKLAWEPPFGILPIIVKQAEBAVAIAKNSDFGLSGSVFTDRF 414
DB 356 KNNLLPLVIDNTYKEMKRVAMBEPPFGILPIITYNSIQEALIDINSQYGLQACIEFTNY 415
QY 415 QKAFADLANKEVEGVTHNNKTGSGPDPNFPGLGKSGAGVQGIKRGISEAMTNVKSIVLD 473
DB 416 ASIEQLALQIESGTININSSSRGPDILPFPGVKGSGFVGQIVDALISMTIKGIIIN 474

RESULT 34
AAE37093 standard; protein; 492 AA.
XX AAE37093;

AC AAE37093;
XX 07-AUG-2003 (first entry)
XX Human 9136 ALDH domain consensus protein.
DE XX

KW Human; aldehyde dehydrogenase; 9136 protein; cardiovascular disorder;
KW cellular proliferative disorder; differentiative disorder; anorectic;
KW endothelial cell disorder; rheumatoid arthritis; arrhythmia; cardiac;
KW heart failure; metabolic disorder; obesity; diabetes; tissue typing;
KW gene therapy; cytostatic; antiinflammatory; cancer; angina; ALDH.
XX Homo sapiens.
XX WO2003033660-A2.
XX 24-APR-2003.
XX 16-OCT-2002; 2002WO-US032971.
XX 16-OCT-2001; 2001US-0329899P.
XX (MILL-) MILLENIUM PHARM INC.
XX Hunter JJ;
XX WPI; 2003-393516/37.
XX Identifying an agent that modulates the level or activity of 9136
XX polypeptide or nucleic acid molecule in a cell, useful for treating
XX cancer, comprises contacting an agent capable of expressing the
XX PT polypeptide or nucleic acid molecule.
XX disclosure; Fig 3, 137pp; English.

CC The invention relates to a member of human aldehyde dehydrogenase (ALDH)
CC family, 9136 protein. The invention also relates to a method for
CC identifying an agent that modulates the level or activity of 9136
CC polypeptide or nucleic acid molecule. The methods, polypeptides and
CC nucleic acid molecules are useful for treating a cellular proliferative
CC or differentiative disorder, e.g. cancer; endothelial cell disorder such
CC as rheumatoid arthritis; cardiovascular disorder such as arrhythmia,
CC angina, heart failure; or metabolic disorder such as obesity or diabetes.
CC The methods are also useful in diagnostic assays, predictive medicine and
CC tissue typing. 9136 DNA is used in gene therapy. The present sequence is
CC human 9136 ALDH domain consensus protein

XX Sequence 492 AA:

Query Match 32.9%; Score 786.5; DB 6; Length 492;
Best Local Similarity 38.6%; Pred. No. 8.9e-57;
Matches 193; Conservative 92; Mismatches 166; Indels 49; Gaps 16;

QY 12 EKKSSVN--QIFILSPIDSS--LGFPVPMATREVDHAKKAGREAL--PMAALTYVERA 65
DB 1 EWDSSASGTFEVDNPNANGVEVIGRVEATADVDVAAPAAKEAFKSGPMMAKVPASERA 60
QY 66 QYLHKAADIIRDEKEIATVLAKEISK--AYNASVTEVVTADLIRYA-----EEG 115
DB 61 RILRIKADIIREREDIALTLTDLCKPLAEAKGDTVEGRALDEIRYVAGMARKMGERR 120
QY 116 I--RLSTSDSGKMDASTGHKLAVIRQPVGIVLAIAPYNYPVNLGSKIAPALIGNV 173
DB 121 VIPLSLATDGD-----ELNVTRRRBPGLGVGVISPMNPPLALMLKLAAPALAGNT 170
QY 174 VNFKPTQGSVSGVLAKAPAEAG--LPAGVFNITTTGSGSELDYIVHEEYVNFNFG 230
DB 171 VVAKPSEQTPPLTALLAELEIBEGANNLPKGVNNVPPGGAEGVQALLSHPIDIKLSFG 230
QY 231 STPVGORIGKLA---GMRPIMLELGGKAGIYLAADADLNAKQIYAGAYDSGORCTAI 287
DB 231 STEVGLIMEAABAANKLKVITIELGSGSPVITFDADADKAVERIVFGAFGNAQVCTAP 290
QY 288 KRVLVVEEVADELAEKISENVAKLS--VEDPPNAT--VTPVIDNSADFIESLVDAROK 344
DB 291 SRLVHESITVDEVEDEKERVKYLKISGPDSDTNIVGPLISEQDFDVLSTIEDGKEE 350
QY 345 GAK-----ELNFKRDKRGLLTPGLFDVHTLDMKLAWEPPFGILPIIRVYVDAEBAVA 396


```

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151308P.
PR 07-SEP-1999; 99US-0151930P.
PR 10-SEP-1999; 99US-0152363P.
PR 13-SEP-1999; 99US-0153707P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155559P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158322P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

```

Query Match 32.8%; Score 784.5; DB 3; Length 293;
 Best Local Similarity 52.4%; Pred. NO. 6.3e-57;
 Matches 150; Conservative 51; Mismatches 84; Indels 1; Gaps 1;

QY 189 LAKAPAEAGLPAAVPTTITRGSEIGDYVEHEBVAFINFTGSTPVGORIGKLAKGRPIM 248

```

DB 1 MVHCFHLAAGPFGKGLISCVTGGKSEIGDFLTMHPGVNVCISFTGG-DGCIATISKAGAVPIQ 59
QY 249 LELGKDAQIVLADADLDNAAQIYAGAYDYGQRTAIKRVLYVEVDELAEKISENV 308
DB 60 MEIGKDAQIVLEADLDLVASANIYVGFSYSGQRTAVKVVLYMESIDAVQOKNATL 119
QY 309 AKLSVDPFDNATVTPVIDNSADFIESIYVVAROKGAELMEFKRDKGLTPGLEDHT 368
DB 120 AKLVGPPEDSDITPVVTESSANFLBGLVMAKKEGATFCQRYRREGNLLPDLIDHYR 179
QY 369 LPMKLAMEBPFGLPIIIRKDAEBAVALANKSDFGIQQSVFTRDPQKADIANKLEVGT 428
DB 180 PMRIMAEERFPGVLPVIRINSVEBGHHCMNSNFGIQQCITRDIINKAILLSDAMEIGT 239
QY 429 VHIINKTGRGPDNFPPLGLKSGAGYQGIIRYIEAMTNYKSVLDM 474
DB 240 VOINSAPARGPDHPFPGQIKSGIGSQGITNSINMTTKVKSIVINL 285

RESULT 36
AAE39889
ID AAE39889 standard; protein; 493 AA.
XX
AC AAE39889;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human aldehyde dehydrogenase domain consensus peptide.
XX
KW Human; cytochrome p450; dehydrogenase/reductase; osteomalacia; psoriasis;
KW lipoxigenase; hydratase; proliferative disorder; hematopoietic disorder;
KW differentiative disorder; carcinoma; sarcoma; leukemia; immune disorder;
KW anorexia nervosa; neural degeneration; muscular disorder; lipid disorder;
KW multiple sclerosis; encephalomyelitis; metabolic disorder; osteoporosis;
KW myasthenia gravis; bone metabolism; immunomodulator; obesity; anorectic;
KW eating disorder; osteodystrophy; arthritis; diabetes; anabolic; rickets;
KW milk fever; diabetes mellitus.
XX
OS Homo sapiens.
XX
PN US2003092658-A1.
XX
PD 15-MAY-2003.
XX
PF 20-JUN-2002; 2002US-00175696.
XX
PR 02-FEB-2001; 2001US-0266140P.
XX
PR 04-FEB-2002; 2002US-00067668.
XX
PA (MEYE/) MEYERS R E.
PA (GLUC/) GLUCKSMANN M A.
PA (RUDO/) RUDOLPH-OWEN L A.
XX
PI Meyers RE, Glucksmann MA, Rudolph-Owen LA;
XX
XX WPI; 2003-765490/72.
XX
XX New cytochrome P450 family members, dehydrogenase/reductase protein.
XX PT lipoxigenase family member and human hydratase useful for treating
XX cellular proliferative and/or differentiative disorders.
XX
XX Disclosure; Fig 16; Opp; English.
XX
XX The invention relates to isolated cytochrome p450 family members such as
XX 33332, 33302 and 32579, dehydrogenase/reductase protein such as 21509,
XX 33770, lipoxigenase family member such as 46638 and human hydratase such
XX as 50090. The invention is useful as diagnostic targets and agents for
XX controlling one or more of cellular proliferative and/or differentiative
XX disorders such as carcinoma, sarcoma or haematopoietic disorders such as
XX leukemia, immune disorders such as diabetes mellitus, arthritis,
XX multiple sclerosis, encephalomyelitis, myasthenia gravis, psoriasis, etc.
XX or metabolic disorders such as obesity, anorexia nervosa, cachexia, lipid

```

[illegible]

XX	US2004132087-A1.
XX	08-JUL-2004.
XX	11-FEB-2004; 2004US-00776871.
XX	31-MAR-2000; 2000US-0193920P.
XX	19-MAY-2000; 2000US-0205675P.
XX	15-JUN-2000; 2000US-0211727P.
XX	02-FEB-2001; 2001US-0266140P.
XX	30-MAR-2001; 2001US-0082390L.
XX	02-APR-2001; 2001US-0010720.
XX	21-MAY-2001; 2001US-00862658.
XX	21-MAY-2001; 2001US-0016380.
XX	15-JUN-2001; 2001US-00882837.
XX	15-JUN-2001; 2001US-0019319.
XX	04-FEB-2002; 2002US-00067668.
XX	20-JUN-2002; 2002US-00175696.
XX	(MILL-) MILLENNIUM PHARM INC.
XX	Meyers RE, Glucksmann MA, Rudolph-Owen LA,
XX	WPI; 2004-533235/51.
XX	New 33312, 33303, 32579, 21509, 33770, 46638, or 50090 nucleic acid
XX	molecules, useful for treating genetic disorders of the membrane
XX	transport, CNS disorders, liver disorders, skeletal muscle disorders, or
XX	cancer.
XX	Disclouseure; SEQ ID NO 21; 248bp; English.
XX	The present invention relates to novel human coding sequences and
XX	proteins for cytochrome P450 proteins 33312 (ADQ48397-ADQ48399), 33303
XX	(ADQ48400-ADQ48402) and 32579 (ADQ48403-ADQ48405); dehydrogenases/
XX	reductases 21509 (ADQ48409-ADQ48411) and 33770 (ADQ48412-ADQ48414);
XX	1ipoxygenase 46638 (ADQ48418-ADQ48420) and hydratase 50090 (ADQ48424-
XX	ADQ48426). The sequences are useful for diagnosing and treating
XX	disorders, such as genetic disorders of the membrane transport
XX	(antioacidurias, cystinosis), CNS disorders (Alzheimer's disease,
XX	epilepsy, Parkinson's disease), liver disorders, skeletal muscle
XX	(cancer), cellular proliferative and/or differentiative disorders
XX	(diabetes), hormonal disorders (diabetes, thyroid disorders), immune and
XX	inflammatory disorders (rheumatoid arthritis, osteoarthritis, ulcer),
XX	cardiovascular disorders (blood vessel disorders, neurophil disorders
XX	(neutropenia, lupus), testicular disorder (mumps) and platelet disorders.
XX	The present sequence is a dehydrogenase consensus sequence used in a
XX	sequence alignment to illustrate the invention.
XX	Sequence 493 AA:
XX	32.5%; Score 777; DB 8; Length 493;
XX	Best Local Similarity 38.5%; Pred. No. 5.5e-56;
XX	Matches 193; Conservative 92; Mismatches 166; Indels 50; Gaps 17
XX	12 EMKGSVN-QEILSPIDDS-LGVPMTEEEVDHAKAGREAL--PAMALLVYERRA 65
XX	1 EWDVSASAKTEVVPANPANGVEVIGVPEATVADVDAAVAKAEAKSGPWWAKVASERA 60
XX	66 QYLHKAAIIERDKKEIATVLAKEISK--ANASATEVVFRTADLIRYAA-----EEG 115
XX	61 RILRLKADLIEREDLEALTLTDIGKRLAERAKGDTVEGRALDEIRRYAGMARKMGERR 120
XX	116 I-RISTSADEGCKMDASTGHKLAVIRROPGVIVLAIAPYVNVNLSGSKIAPALLIGNV 173
XX	121 VIPSIATDGR-----ELNVTREPIGVGVISPMNFPILLALMKLAPALAAAGNT 170
XX	174 VMFEPPTGVSGLVLAAPAEAG--LPAGVNTITRSGEIGIYIYEHHEVNFINTFG 230
XX	171 VLVKPESEPTPLTALLLELIEAGANNIPKVVVNVVPFGAEVGCALLSHPIDISFTG 230
XX	231 STPGGORIGKIA---GMRPIMLEIGSGRAGIVLADADIDNNAKQIVAGAYDVSQGRCTAI 287

OS Bacteria.
 XX US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS.
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 PS Claim 1; SEQ ID NO 1073; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX
 SQ Sequence 463 AA;

Query Match 31.7%; Score 759.5; DB 8; Length 463;
 Best Local Similarity 37.1%; Pred. No. 1.5e-54;
 Matches 175; Conservative 95; Mismatches 183; Indels 19; Gaps 8;

QY 8 YNNGEWSKSVNO--IRILSPIDSSLGFPAMTRREVDHAKAGRALPAMAAALTYERA 65
 DB 2 FIDGKM--INREMDVINPYSLEVIKIKIPALSREBAKAIPTAEYKVKVMKNLPITKXY 58
 QY 66 QYLHRAADIERDKERIEATVLAKEISKAYNAVSTEVRTADILRYAAEGIRLSTADRG 125
 DB 59 NTLMMNIAKIKERKEBELAKIALIDAGKPIKQARVEVERSIGFPKLA--FVYKERDREV 115
 QY 126 GKMDASTGKLAIVIRROPYGIYLAIPVYPNVNLSSKIALPALIGNVVMFKPPQGSVS 185
 DB 116 IPSD---DRLIFTRREPVGIVAIPTPFNPPLNSHKIAPAIATGNVIVHPPSSKAPLV 171
 QY 186 GLVIAK---AFPAEAGLPAGVPTITNGRSEIGDYVHEHEVNFITNSTPVGQIGLU 241
 DB 172 CIELAKIENALKKNVPLGVNLLTAGEVGVDEIVNEKVMNLSFTSSKVGELITTK 231

QY 242 AGMRPIMLEGGKADAGIVLADADLDNNAKOIVAGAYDYSGORCTAIKRVLVVEEVADELA 301
 DB 232 AGFKKIALLEGVNPVNITLXODALINKAVNALIKGSFTIYAGQVCSVGMILVDESIADEFI 291
 QY 302 EKISENVAKLSVGDPEFNAT-VTPYIDNSADFLFSLVYDARQKAKELNFKKDGRLTT 360
 DB 292 EMFVNKAKVLNVGNPLDEKTVGPLISVEHAWEKVEKAIIDEGKLLGGKRKALFY 351
 QY 361 PGLPDPHTLDMKLAEEBPPGPIPLPIRYKDAEBAVALANKSDPGLQSSVFTRDFOKARDI 420
 DB 352 PTLIE-VDRNILLCTETTFAPVPIIRINE-BEMIDINSTEYGLHSAIFINDIKSLKF 409
 QY 421 ANKLEVGTVHINNKTGRGPDNPPFLGKSGAGVQGRYSIEMATNVKSIYL 472
 DB 410 AENLEFGGVVINDSSLFPQDNMPPFGVYKSGIGREGVXYAMEMSNIKTIIL 461

RESULT 40
 ADN18535
 ID ADN18535 standard; protein; 455 AA.
 XX
 AC ADN18535;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DS Bacterial polypeptide #1188.

XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 1188; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,

	Query Match	30.8%	Score 737;	DB 8;	Length 476;
	Best Local Similarity	35.0%;	Pred.No.1,2e-52;		
	Match 168;	Conservative	96;	Mismatches 200;	Indels 16; Gaps 6
Oy	4	EYQNVNGEWSKSVNOIEILSPIDSSLGFPVPMTEEREVDHAKAGREALPAWALTVEE	63		
Dd	3	EKKMLIGSEWVNDKEETIVITYPYTRPIGRVRGDEKDVKAEIRAKEGFKEIFSLTAYE	62		
Oy	64	RAQYLHKAADIERDKEEIATVLAKESIKYNNASVTEVRTADLIRYAABEGIRLS----	119		
Dd	63	RHILMRAPAAQLTKERAEPEAKTLVLEVGKTIREARTREVORALQTLLFSAEARKVNGETF	122		
Oy	120	-TSADEGGMDASTGHKLAVIRROPGIYLAIPNYPYNLGSCKIAPPLIGNVMFMP	178		
Dd	123	PIDAHPNG-----GKLGYFRVPVGIYSATTPFNPLNSHKAPALPAGNAVALTKP	176		
Oy	179	PTQGSVSGIVLAKAFEAELPAGVFNTITGRSGEISDIYVEHEEVFIINTGSTPVGORI	238		
Dd	177	SERTFLTPLMGEILLLEAGVPPRKALSVIRGYG-DVGKAMTHPDVAVVSFTSRRKGDI	235		
Oy	239	GKLAGRPRIMTELGGKDAGITVLADADDLNAAKOIYAGAYDSGORCTAIRVLVVEVAD	298		

RESULT 44
 ADS28216
 ID ADS28216 standard; protein; 482 AA.
 XX
 AC ADS28216;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #17249.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 17249; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 482 AA;

Query Match 30.7%; Score 735; DB 8; Length 482;
Best Local Similarity 35.0%; Pred. No. 1.7e-52;
Matches 164; Conservative 90; Mismatches 209; Indels 6; Gaps 3;

QY 9 VNGEMKSSVNOIEILSPIDSSIGFVPAMTREEDVHAMKAGREALPAMALTYVEAOYL 68
DB 12 ISDQOHGQOSIEVIDPGTNHVIAIVPAPATEADAKAIEAKGEMARQMIHERMALI 71
QY 69 HKAADIIEEDKEEIAIVLAKESKAYNAVSTEVTRADILRYAEEGIRLSTSDGCKM 128
DB 72 KTAIEIVTRDELFAOTIAKESKTRRARSSEVRRICETILTSGEAKKLHGSETTIPFSQ 131
QY 129 DASTGH--KLAIVIRROPVGIIVLAIAPYNPVNISSGKIAPALIGNVNFKPPTQGSVG 186
DB 132 ---PGERRVGVFYFRPPIGIIIVATFPNDPLNVAKHIGPAAIAGSLIIKSSFTPLSA 188
QY 187 LVLAAPFAAGPAGVFNITIGRSGEIGYIVHEBEVNFINTGSPVQORIGKLAGMP 246
DB 189 LELVLALEAGLPPKIKVQVITGHGVIPTLTGHRVRLISFTGSGYETEKLARSAGVKK 248
QY 247 IMLEIGKRGAGIVLADADLDMNAKQIVAGAYDYSGORCTAIRKVLVEEVADELAKISE 306
DB 249 IAMEIGSNPTIYLDALMEAVASCVAFCAGANCICGVQIITYEGSVNFSFISFA 308
QY 307 NVAKLSVGDPEFDMAT--VTPVIDNSADFTESLVVDAROKAKELNEFKKDGRLTPELPD 365
DB 309 QTRQLRLDGKQSEBETDIGMISEKAKRIERWVEBAKEGAVLTGRRGTAVFEPTVLT 368
QY 366 HTYLDKLAWEPPFGILPIIRYKAEBAVAIANKSDPGLQSSVFTPRDQKAPDIANKLE 425
DB 369 NVSPSSRLAKEBAFAIVIEGVHSLTEAIAARANDVDFGLQGLFTNNLTAAFSALIEKLE 428
QY 426 VGTTHINNTKGRGPNPFLGLKSGAGVQGIYSTIEMTNVYSITLDM 474
DB 429 VGGIWNDSDVRIIDAMPFGIIGKSGLGREGVRAIEEMTEOKVAFHL 477

RESULT 45

ADM26502
ID ADM26502 standard; protein; 471 AA.

XX ADM26502;

XX 20-MAY-2004 (first entry)

XX Hyperthermophile Methanopyrus kandleri protein #1108.

XX hyperthermophile; protein stability enhancement;
XX protein activity enhancement.

XX Methanopyrus kandleri.

XX WO2003076575-A2.

XX 18-SEP-2003.

XX 04-MAR-2003; 2003WO-US006664.

XX 04-MAR-2002; 2002US-0361742P.

XX 14-MAY-2002; 2002US-0380423P.

PR 16-SEP-2002; 2002US-0410974P.

(FIDE-) FIDELITY SYSTEMS INC.

PA (MALY/) MALYKH A.

PI Slesarev AI, Pavlov A, Pavlova N, Kozayavkin S;

DR WPI; 2003-748383/70.

DR N-PSDB; ADM27081.

PT New isolated nucleic acids encoding any of about 1700 Methanopyrus
PT kandleri proteins, and the encoded proteins, useful as a medicaments or
PT as diagnostic agents.

PS Claim 31; SEQ ID NO 1108; 1023bp; English.

CC The invention comprises the amino acid sequence of proteins from the
CC hyperthermophile Methanopyrus kandleri, the invention also comprises the
CC complete genome from Methanopyrus kandleri. The Methanopyrus kandleri
CC proteins of the invention are useful for enhancing the stability and/or
CC activity of other proteins. The Methanopyrus kandleri genome is useful in
CC a variety of diagnostic and analytical methods. The present amino acid
CC sequence represents a Methanopyrus kandleri protein of the invention.

XX Sequence 471 AA;

Query Match 30.7%; Score 734; DB 7; Length 471;
Best Local Similarity 36.7%; Pred. No. 2e-52;
Matches 173; Conservative 90; Mismatches 201; Indels 8; Gaps 5;

QY 4 EYQNVNGEMKSSVNOIEILSPIDSSIGFVPAMTREEDVHAMKAGREALPAMALTYVE 63
DB 2 EYQCLIGEMLEGDREIVENPDGSEVGRVLA--PEVDEVALLDKAREGQRMRERPTYE 60
QY 64 RAQYHKAADIIEERKEEIAIVLAKESKAYNAVSTEVTRADILRYAEEGIRLSTSD 123
DB 61 VREALAEAAHLKKEDELAELILEGGKPIRDAREYVRRREVLRISAARERLYGTEL 120
QY 124 EGGKMDASTGH--KLAIVIRROPVGIIVLAIAPYNPVNISSGKIAPALIGNVNFKPPTQ 181
DB 121 PG---DAQRGRTABILLVREBPVGLSTIYNPFLLPHTKLGPALAAKGSVNHKPAIV 177
QY 182 GSVSGVLAKAPAEAGLPAVFNITIGRSGEIGYIVHEBEVNFINTGSPVQORIGKL 241
DB 178 TPLSSRLAELILDAGVEPLALQYVVGEGALGEBELA--RADFDALSFGRSRSVGBHLREI 236
QY 242 AGMRIMLEIGKRGAGIVLADADLDMNAKQIVAGAYDYSGORCTAIRKVLVEEVADELA 301
DB 237 SPIPRITTELGNDVIVDETADEAAAEAAVRCGAAAGOVCIATERAIVVEDVTEBFL 296
QY 302 EKISENVAKLSVGDPEFDMAT--VTPVIDNSADFTESLVVDAROKAKELNEFKKDGRLT 360
DB 297 EAAVVAASLKVGDPLDEDTDVGFLIDGAVEKARRHVEDAVERGAELITGGEPEGRFLP 356
QY 361 PGLFDHTYLDKLAWEPPFGILPIIRYKAEBAVAIANKSDPGLQSSVFTPRDQKAPDI 420
DB 357 PTVLADVPEDALVAREETFPGLVPIYRAKDPEEAIRINSTDYGLHAAVFTERLDBAVKA 416
QY 421 ANKLEVGIVTHINNTKGRGPNPFLGLKSGAGVQGIYSTIEMTNVYSITL 472
DB 417 TRKLEAGGIVIVESTIYRADYMPFGVAGSAGVGRGVQAVRFEETKTVVI 468

RESULT 46

ABO78556
ID ABO78556 standard; protein; 526 AA.

XX ABO78556;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #10731.

ID ABU28726 standard; protein; 482 AA.
 AC ABU28726;
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #14253.
 XX
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Escherichia coli.
 XX
 PN MO200277183-A2.
 XX
 PD 03-OCT-2002.
 PF 21-MAR-2002; 2002MO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362659P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haasebeck R, Ohlsen KU, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
 XX
 DR N-PSDB; ACA32596.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 56650; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 482 AA;
 Query Match 30.3%; Score 726; DB 6; Length 482;

Best Local Similarity 34.3%; Pred. No. 9,8e-52;
 Matches 164; Conservative 92; Mismatches 206; Indels 16; Gaps 7;
 QY 6 QNYVNGEWSKSVN--QIEILSPIDSSIGFVPAMTRREVDHAMKRGALPMWALTVE 63
 DB 12 QALINGEMDANNGEAIDVTNPANGDKIGSVKMGADETRAAIDANRALPMWALTVE 71
 QY 64 RAQYLHKADITIERDKERIATYLAKEISKANVASTVEVTRDILRYAEEGRISTSD 123
 DB 72 RATILNMFNLMMEHQDDILRLMTLEQKPLAEAGEISYVASFIEWFAEEGRITYGDTI 131
 QY 124 EGGKMDASTGHTAVIRQPGIVLAAPYNPVNLGSGKIPALIGVNVMEKPTQGS 183
 DB 132 PGHQAD----KRLVIKQPIGVTAITPWNPPAMITRKAGPALLAGCTMVKPSQTP 186
 QY 184 VSGLVLAKEAPAGAGPAGVFTNTTIGSGSEIGDYIYEHEVNFINTGSTPVGQRTGKLAG 243
 DB 187 FSALMLAEALIRAGVPAVFVNTGSAVGNELSNPLVRKLSFGSTIEGRQLMEQCA 246
 QY 244 --MRPTMEELGSKAGIVLADLDLNAKQIVAGAYDSGRCCTAIKRYLVVEVADDELA 301
 DB 247 KDIKKVSEELGNNAPFIVPDDADLDKAVGALASKFRNAGQTCVCANRLYVDGVDYDRA 306
 QY 302 EKISENVAKLSVGDPEPN--ATVPTVIDNSADFISLVYDAROKAKEL--NEPKRDR 357
 DB 307 EKLQQAVSKLHTIGDLDNGVTTIGPLIDKAVAKVEEHIALDEKARVYCGKAKHERGDN 366
 QY 358 LITPGLFHDVTLDMKLAWEPPGPILPIIRVDAEAAVAIANKSPFGLOSVFTDFQKA 417
 DB 367 PQPFTLVDPVANAIVKSKETFGPLALPFRFDEADVIAQANDTEGGLAAVYARLDSRV 426
 QY 418 FDIANKLEVGVTHINNKTG--RQPDNPPFLGKSGAGVCGIGYISFAMNWKSYIYLD 474
 DB 427 FVGEALFEGVIGIN--TGIISEVAPFGGIRASGIGREGSKYGIEDYELIKYMGIGL 482
 RESULT 50
 ID ADS45081
 AC ADS45081 standard; protein; 482 AA.
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #23511.
 XX
 KM Recombinant DNA construct; transformed plant; improved plant property;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
 KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KM bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX Claim 1, SEQ ID NO 23511; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transforming plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 482 AA;

Query Match 30.3%; Score 726; DB 8; Length 482;

Best Local Similarity 34.3%; Pred. No. 9, 8e-52; Matches 164; Conservative 92; Mismatches 206; Indels 16; Gaps 7;

6 QNVNNGEMSSVNV-QIETLSPIDDSLGFPVPMTRTEEDVHMKAGREALPMAALTYVE 63
12 QALINGEWLDANNGEAIVDTNPANGKLSVPMGADETRAAIDANRLPAMRALTAKE 71
64 RAQYLHKAADIIEERKEEIAVLAKESKAYNAVTEVVRADLRVAAEGRIRLSTSD 123
72 RATILFNWNLMMHEODDLARLMTLEQGRPLAEAKGEISVYASFIEWFBEGRIRVGDRT 131
124 EGGKNDASIGHKLAIVRRQPVGIVALIAPYNPVNLSSGIAPALIGVNVWFKPTQGS 183
132 PEGQAD----KRLIVIKQPIGVTAITPWNPPAAMITRKAGPALAAGCTWVLKPSQTP 186
184 VSGVLVAKAFAEAGLPAGVFNTITGSGEIGDIYVHEEYVNFNFTGSPVQORIGKLAG 243
187 F8ALALAEALIRAGVAGVFNVTGSAAGVNELTNPVKLSFGSTIEIGQLHEQCA 246
244 --MRPIMELGKGDAGIVLADADLDNAKQIVAGAYDSGORCTAIKRLVVEEVADELA 301
247 KOIKKVSLELGGNAPITFDDADLDKAVGALASKRRNNGQTCVCANRLVYDGDVDRFA 306
302 EKISNVAAKLSVGDPPDN-ATYTPVUIDNSADFIESLVDAROKAKEL---NEFKRGR 357
307 EKLQQAVSKLHTIGDGNVITIGPLIDERAIVAKVEEHIDALEKAGARVCGGKAHRGCGN 366
358 LITPGFHDVITLDMKLAWEPRGGLPIIRVQDAEBAVAIANKSDGLOSVFTPRFOKA 417
367 FPGPITLVDVPAAKVSKSEETFGPLALFRFDEADAVIAOANDTEGGLAAVFAARLSRV 426
418 FDIANLLEVTGVIHNNKTG-RGPDNPPLGLKSGSAGVQIRYSIEMTNVKSIVLDM 474
427 FAVGEALVEYIGVIN--TGIISNEVAPFGIKASGIGRGSSKXGIDYIEIKTMCIQL 482

RESULT 51
ADN46591

ID ADN46591 standard; protein; 511 AA.

XX ADN46591;

XX 01-JUL-2004 (first entry)

XX Thermococcus kodakarensis KOD1 protein sequence Segid469.

XX gene disruption; gene targeting; marker gene; transformation;

XX homologous recombination; hyperthermostable archaeobacterium; KOD1;

XX forensic science; food; drug inspection; molecular biology; immunology.

XX Thermococcus kodakarensis.

XX MO2004022736-A1.

XX 18-MAR-2004.

XX 29-AUG-2003; 2003MO-IB003597.

XX 30-AUG-2002; 2002JP-00319011.

XX (NISC-) JAPAN SCT & TECHNOLOGY CORP.

XX Imanaka T, Atomi H;

XX WPI; 2004-257583/24.

XX Method for disrupting targeted gene in genome of organism particularly

XX PT thermostable bacterium and with genome chips for analysis, applicable in

XX PT studying gene structure and functions.

XX Claim 9; SEQ ID NO 469; 598bp; Japanese.

CC This invention relates to a novel method for targeting disruption of an
CC arbitrary gene in a genome of an organism which comprises providing the
CC whole sequential data of the genome of such organism, selecting at least
CC 1 arbitrary region in the sequence, providing a vector that contains a
CC sequence homologous with the selected region and a marker gene,
CC transformation, and homologous recombination. The genome is preferably
CC the genome of a hyperthermostable archaeobacterium, particularly
CC Thermococcus kodakarensis KOD1. The method is for targeting the
CC disruption of a gene in the genome of an organism, which is applicable in
CC studying gene structure and functions as well as enzyme activities of
CC encoded proteins and useful in medicine, forensic science, food or drug
CC inspection, molecular biology and immunology. With this method, the
CC disruption of a gene at an arbitrary position in a genome can be achieved
CC efficiently and reliably. The present sequence is that of a protein
CC encoded by the genome of Thermococcus kodakarensis which was derived
CC using the method of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 511 AA;

Query Match 30.2%; Score 722; DB 8; Length 511;

Best Local Similarity 35.0%; Pred. No. 2, 3e-51; Matches 168; Conservative 99; Mismatches 193; Indels 20; Gaps 7;

4 EYQNVNNGEMSSVNVQIETLSPIDDSLGFPV---AMTRBEVDHMKAGREAL---PAM 56
28 EFAITVNGEWYTGTAIEVRSFIDSLIARVSLSMALSNRAVAAVYAGREHIDTGG- 86
57 AALTYERAQYLHKAADIIEERKEEIAVLAKESKAYNAVTEVVRADLRVAAEERI 116
87 -----EKLAEALFKVAELIRBSFDDFTALVLDACKPISNARGVETATIERLEKTTMEFG 141
117 RLSTADGEGKNDASIGHKLAIVRRQPVGIVALIAPYNPVNLSSGIAPALIGVNVWF 176
142 RLIGDIYIGDWSAELSGE-GIVKREPGVLAISPPYPLFIISTAKIVPALAGNAVL 200

QY 177 KPPTGSGVGLVAKAFKAEAGLPAGVFNITITGRGSEIGDYIVHEEVPNIFNPGSVPGO 236
 DB 201 KRSQDPPLPPLLSRVLQALGIPESVYHLLTVPGA-LMDSILADRIRAVFTGSTEVEAS 259
 QY 237 RIQKLAGMPIMLELGGKDGIVLADADLDNNAKOIVAGAYDYSGORCTAIRKVLVEEV 296
 DB 260 HILSMGIGIFVHMEIGKDPVAVLDDAPLEETVEKLVKGVNSVSGRCARILIAEEGI 319
 QY 297 ABLEAKISENNAKLSVGPFP--DNATVTPVIDDNSADPISLVDAROKAKELNEFGR 354
 DB 320 YSQLKRELVAALSKLEPENPLEDEDIMGPIINERSAKIEEYVRDALKEKGVPLTGFGR 379
 QY 355 DGRLLTPGLED--HYTLDMKLAWEPPFGPIPIIRVKDAEBAVALANKSDGELSSVPT 411
 DB 380 KQAYVPVLLERSREVLPGIRAFQEDVFGPLTILVSNDEDAVELANSSRGLDAVPS 439
 QY 412 RDFOKAFDIANKLEVGTVHINNKTGRGPNFPPLGLKSGAGVQGIYSIEAMTNVKSIV 471
 DB 440 GDDSRARVAVARLEVGAVFINERPRGIGYVPRGKMGDSIGREIGISIELITTTKTIV 499

RESULT 52

ABU19008
 ID ABU19008 standard; protein; 483 AA.

AC ABU19008;

DT 19-JUN-2003 (first entry)

DE Protein encoded by prokaryotic essential gene #4535.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Bacillus anthracis.

OS WO200277183-A2.

PN 03-OCT-2002.

PD 21-MAR-2002; 2002WO-US009107.

PF 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haebelbeck R, Ohlsen KU, Zyskind WJ,

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-029926/02.

XX N-PSDB; ACA22878.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 46932; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 483 AA;

Query Match 30.2%; Score 721.5; DB 6; Length 483;

Best Local Similarity 35.7%; Pred. No. 2,3e-51; Mismatches 171; Conservative 82; Mismatches 203; Indels 23; Gaps 8;

QY 8 YVNGEMKSSVNOIEILSPIDSSLGFPVAMTREEDVHAMKAGREALPAAALTYTERAY 67
 DB 16 YINGEMITLQEOIEVNNPATKEIFATVPKGVTEAKQAVDAHEAFKMSKLTADRAAK 75
 QY 68 LHKADIIIRDBKEEIAITLAKETISKAYNASVEVYRTADLIYYAEBGRISTADEGK 127
 DB 76 LKKWFTLIDENKEEIAITMTEQGRFAEALGEVYVANSFVWYABEGRV-----YGEH 130
 QY 128 MDASTGHKLAIVRRQVGVILAIAPYVNVNLSGSKIAIPALGCVNVMFKPTGSGVSG 187
 DB 131 IPASHPNKRIILVMKQPVGMALITPNWFPAMITTKVAPALAACTAAVVKPSQPTPL 190
 QY 188 VLAKAFKAEAGLPAGVFNITITGRGSEIGDYIVHEEVPNIFNPGSVPGO-NR 245
 DB 191 KLABLAHEADIPKGYINIVTGSAKAIADTWMDGVRKVSFTGSEIGELMAASAQTWK 250
 QY 246 PIMLELGGKDGIVLADADLDNNAKOIVAGAYDYSGORCTAIRKVLVEEVADELAEK 305
 DB 251 KVSLELGGHAPETVWMDADLDRAVAIVGSKPRNAGQCTICTNRVFEVEVEVAEVEKE 310
 QY 306 ENVAKLGVGDPF-DNATVTPVIDDNSADPISLVDAROKAKELNEFGR---DGRLLT 360
 DB 311 KAVGOLKVGDDGFGDGTGVPILIDENAVSKVQDHIEDALOKGTVLYGGQKVAELDGHFMQ 370
 QY 361 P---GLFPHVTLDMLKLAWEPPGPIPIIRVKDAEBAVALANKSDFGLOSVFTRDPK 417
 DB 371 PVIIGLANDTMLCMN---EETFGVAPVAKFTVEVERANHTYEGLAAYIFTDISGA 427
 QY 418 PDIANKLEVGTVHINNKTGRGPD--NPPLGLKSGAGVQGIYSIEAMTNVKSIVLDM 474
 DB 428 FOISALBEYGIILGNDGL---PSVAQAPFGPKESGIGREGHFGIEBYLEIKYISLGL 483

RESULT 53

AAAR55135
 ID AAR55135 standard; protein; 488 AA.

AC AAR55135;

DT 16-OCT-2003 (revised)

DT 18-JAN-1995 (first entry)

XX Heat resistant aldehyde dehydrogenase.

DB aldehyde dehydrogenase; heat resistant; alcohol production;

KM Bacillus stearothermophilus.

```

OS Geobacillus stearothermophilus.
XX Key Location/Qualifiers
XX CDS 1266..2732
XX /*tag= a
XX /product= "heat resistant aldehyde dehydrogenase"
XX
XX JP06113839-A.
XX
XX 26-APR-1994.
XX
XX 30-MAR-1992; 92JP-00074631.
XX
XX 30-MAR-1992; 92JP-00074631.
XX
XX (MARU-) MARUKAN SU KK.
XX
XX MPI: 1994-172737/21.
XX N-PSDB; AA065587.
XX
XX Heat-resistant aldehyde dehydrogenase from Bacillus stearothermophilus -
XX and corresp. DNA, useful for alcohol prodn.
XX
XX Claim 1; Page 8-10; 11pp; Japanese.
XX
XX This sequence shows an aldehyde dehydrogenase (ALDH) (see also AA065587),
XX that is heat resistant from Bacillus stearothermophilus and requires no
XX acetyl CoA. The fused protein of ALDH with an alcohol dehydrogenase
XX derived from a transformant which contains both structural genes can be
XX used in production of alcohol. (Updated on 16-OCT-2003 to standardise OS
XX field)
XX
XX Sequence 488 AA;
SQ
Query Match 30.2%; Score 721.5; DB 2; Length 488;
Best Local Similarity 34.4%; Pred. No. 2,4e-51;
Matches 166; Conservative 104; Mismatches 192; Indels 21; Gaps 9;
QY 3 KEYQYVNGEMKSV-NOIE-ILSPIDDSL-GVPVAMTREVDHAMKAREALPAMAL 59
DB 8 KYFFVYINGNMWTSVSNVPEPSINPANRHDIVGYGRSLDEYVNEAVTANERQJSMWR 67
QY 60 TTY#QAQYIAKADIIERDKETATVLAKEISKAYASVTEVVRTADLIRYAEGSIRUS 119
DB 68 SCVENGEVLYKAAHLEQCLDIAETMTREMGKTLAEAKAETMRGVHILRYAEGGAR-- 125
QY 120 TSADSGKMDASTGHKLAVIRROPVGIYALAPYVAVVLSGSKTAPALIGAVVWFKRP 179
DB 126 ---KIGDVISSDSGLFTTTPVPSGVVISPMNPVAPLPTWKAPALVIGTVVLPKA 182
QY 180 TQGSVSGVLAKAFAPAGVFTTITGRGSEIGDYIEHSEVNFINFSTPGVQORIG 239
DB 183 SFTATVAKVIECFHEAGFPKGVVNMVVGSGSVGGGINHDDIVFTTGSNTYKQNG 242
QY 240 KLAAGR--PIMELGSKAGIYLADADLNNAKQIYAGAYDYSQGRCTAIKRYLVEEVA 297
DB 243 RAAPFRGARYQLEMGGKMPVIYAKDADLDAVEGTISGGLRSTGQCTATSVFTIEREYV 302
QY 298 DLAKISENVAKLSGVDPFDMAT--VTPVIDNSADFIESIYVVDARQKAKELNEFKR-- 354
DB 303 EFPKAKLIERVKQLKINGLDAETWGPASBSQFTTVAISYLEKGSBEAKLIYGGNRCL 362
QY 355 -----DGRLLTPGLPDHVTLLDMKLAMEEPFGPLPIIRYKDAEVAVALANKSDPGLQSSV 409
DB 363 EGEIANGPFVEPTIFEDVDLQMTIAREEIFGPVLAIIQVDSIEBAIKLANDREYIGLSASI 422
QY 410 FTRDQKAPDIANKLEVGTVHINNTGGRPDNFPPLGLK--GSGAGVQIGIRYSIEAMTV 467
DB 423 YKTNIGMLEFTIKDIEAGLIKVNATAGYEQAPFGGMKROSSSHSREOG--QAIIEFTSI 481
QY 468 KSI 470
DB 482 KTV 484

```

```

RESULT 54
ABB47807
ID ABB47807 standard; protein; 488 AA.
XX
XX ABB47807;
XX
XX 05-FEB-2002 (first entry)
XX
XX Listeria monocytogenes protein #511.
XX
XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX vitamin B12; bacterial infection; disease.
XX
XX Listeria monocytogenes.
XX
XX WO200177335-A2.
XX
XX 18-OCT-2001.
XX
XX 11-APR-2001; 2001WO-FR001118.
XX
XX 11-APR-2000; 2000FR-00004629.
XX
XX (INSP) INST PASTEUR.
XX
XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P,
XX Duesberg O, Chetoui F, Nedjari H, Glaser P, Kunst F, Coessart P,
XX Daniels U, Geobel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA,
XX Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A,
XX Chakraborty T, Domann B, Hain T, Berche P, Charbit A, Durant L,
XX Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
XX Madueno E, De Pablo B, Wehlend U, Kaerst U, Entian K, Hauf U;
XX Rose M, Voss H;
XX
XX MPI: 2002-010914/01.
XX
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX and prevention of Listeria and related bacterial infections, and related
XX polypeptides.
XX
XX Claim 6; SEQ ID NO 512; 192pp; French.
XX
XX The present invention relates to the genome sequence of Listeria
XX monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
XX it are useful for selecting probes and primers for detecting genes in L.
XX monocytogenes and related organisms, and for studying genetic
XX polymorphisms and other genomes. The present sequence is a protein
XX encoded by the genome sequence of the present invention. Proteins
XX expressed from the genome sequence are useful for raising specific
XX antibodies, identification of L. monocytogenes and related organisms, and
XX for biosynthesis and biodegradation, especially biosynthesis of Vitamin
XX B12. The genome sequence and proteins encoded by it are also useful for
XX selecting compounds that regulate gene expression and cell replication
XX and modulate L. monocytogenes-related diseases. In addition, the genome
XX sequence and proteins encoded by it are useful in pharmaceutical and
XX vaccines compositions for the treatment or prevention of infections by L.
XX monocytogenes and related organisms. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 488 AA;
SQ
Query Match 30.0%; Score 718; DB 5; Length 488;
Best Local Similarity 33.8%; Pred. No. 4.6e-51;
Matches 161; Conservative 96; Mismatches 201; Indels 18; Gaps 6;
QY 8 YVNGEMKSVNO--TEIUSPIDDSLGVPVAMTREVDHAMKAREALPAMALTVVRA 65
DB 16 FINGKMTGDNKTKETQIVNPANGVYAKIAQAGSBETKAKIAKADAPPDMAKMLADRV 75

```

Qy	66	QYUHQADADIEEDKEIAIVLAKESKAINASTEVVRADLIRYAASEBIRLSTADEG	125
Db	76	KLHHTADIMMEKADTLAKIMTLBOGPKLESKEGLVTEVNERFAAEARRL----	YG 130
Qy	126	GMDASTGHKLAVIRROPQVILATAPVYVNVLSGSKTAPALIGGVVWFKPPTQOSVS	185
Db	131	ETIPAPNNHAF-IYKKQPIGVAAITPPMFPGGMVTRKLAIPALATGNTIYLKSSGDTPLS	189
Qy	186	GLVLAKAFEAAGLPAGVFNITTTGRGSEIGDYIYEHEVNFINEFTSGTPVGQIRIGLAG--	243
Db	190	ALATFEIPEEAGLPKGVANIWMGSSKEIGBTLTSDSDVRLTFTGTSGTVGQTLFKQSADT	249
Qy	244	MRPIMLELGGKQAGIYLDADLDNNAKOIYAGAYDYSGQRCTAIKRVLYVEEYVADELAEK	303
Db	250	LKKIIELELGGHAPFIYFVDANLDAAVNDVAAKFRNNQGVCCSPNNIIPVAKETKEFTKA	309
Qy	304	ISENVAKLSVBPBDVATVTPVIDDVSADPISLSJVDAORKGAKELNEFR-----DG	356
Db	310	LVAKYBOLKVGNGLDVNVGPIRLIREADIKIDIKOLKNAITEKAKVLTTGGRLTGSDDYDKG	369
Qy	357	RLUTRGELFHDVTLDMKLAEMEBPFGILPIIRIVDAEAAVAALANKSDFGLOSSFTRDFOK	416
Db	370	NFYKPTVLNDVTRKODIFYEETFGPVITPITFETEDBAIEMANDSEBGLASYPYTDIAR	429
Qy	417	AFDIAKLEVGIVYHINKTGRGDNFPFJGLKSGSGAGVOCIRISIEAMTVKSYVL	472
Db	430	VEKVAAALEYGVNVAINEIASNPET--PFGGVKHSIGSGRENGHGMEEBYIYQVKINL	484

RESULT 55
 ID ABU32914 standard; protein, 488 AA.
 XX ABU32914;
 XX
 AC
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #18441.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS *Listeria monocytogenes*.
 XX
 PN MO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342821P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA36784.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 60838; 1766bp; English.
 CC
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid

encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
 the target prokaryotic essential genes. Note: The sequence data for this
 patent did not form part of the printed specification, but was obtained
 in electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

[illegible]

02-DEC-2004 (first entry)
Bacterial polypeptide #18199.
Recombinant DNA construct; transformed plant; improved plant property; cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic; pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbohydrate; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polypeptide.
Bacteria.
US2003233675-A1.
18-DEC-2003.
20-FEB-2003; 2003US-00369493.
21-FEB-2002; 2002US-0360039P.
(CAOY/) CAO Y.
(HINK/) HINKLE G J.
(SLAT/) SLATER S C.
(CHEN/) CHEN X.
(GOLD/) GOLDMAN B S.
Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
WPI; 2004-061375/06.
New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.
Claim 1; SEQ ID NO 18199; 122pb; English.
The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polypeptide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

Query Match	29.9%	Score 716.5;	DB 8;	Length 493;
Best Local Similarity	34.6%	Pred. NO. 6.3e-51;		
Matches 167; Conservative	97;	Mismatches 191;		
				Indels 27; Gaps 9;

Oy 8 YVNGEN--KSSVNQIEILSPIDDSLGFPVPAMTREEDVDHAKAGEALPAWAALTVERA 65
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Dd 6 YIDGQWNNSSSGKTVDKKSPTVGIVGREAAATDDVDRAIDAEDAFAWANDLGSVERS 65

[illegible]

RESULT	57
ADSD29560	
ID	ADSD29560 standard; protein; 480 AA.
XX	
AC	ADSD29560;
XX	
DT	02-DEC-2004 (first entry)
XX	
DE	Bacterial polypeptide #18593.
XX	
KW	Recombinant DNA construct; transformed plant; improved plant property;
KW	cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW	pathogen tolerance; pest tolerance; plant disease resistance;
KW	cell cycle pathway modification; plant growth regulator;
KW	homologous recombination; seed oil yield; protein yield; carbohydrate;
KW	nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KX	bacterial polypeptide.
XX	
OS	Bacteria.
XX	
PN	US2003233675-A1.
XX	
PD	18-DEC-2003.
XX	
PF	20-FEB-2003; 2003US-0036993.
XX	
PR	21-FEB-2002; 2002US-0360039P.
XX	
PA	(CAOY/) CAO Y.
PA	(HINK/) HINKLE G J.
PA	(SLAT/) SLATER S C.
PA	(CHEN/) CHEN X.
PA	(GOLD/) GOLDMAN B S.
XX	
PI	Gao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX	
DR	WPI; 2004-061375/06.
XX	
PT	New recombinant DNA construct comprising a promoter positioned to provide
PT	for expression of a polynucleotide encoding a polypeptide from a


```
QY      6 QNYVJGEMKSSVN--QIELTSPIDDSLSGFVPMTRBEEVDHAKKAGREALPMAALTVE 63
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      12 QALLIGOMDADNBDVIATVPNPNANGEOILSVPPMGADDEFREALEANRRALPARRALTAKE 71
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      64 RAQYLHKADIIEERDKBEIATVLAKEISKAYNASVEVVRTADLIRYAABEGTRLSTAD 123
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      72 RANTLRMPDLWMENODDLARLMTLTLEQGKPLAEKKEISVAASFIEFPAEBEGRRIYDPT 131
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      124 EGGKRDASTGHKLAVITRROPVGIVLAIAPIYRVNLSSGIAPALIGENVVMFKPPTQGS 183
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      132 PGHOAD----RRLVITIKOPIGVTAITWNPFPAAAITRKXGAPALAACTMVLKPASQTP 186
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      184 VSGLVIAKAFAEAGLPAGVENTITGKSIBGDYIVHEBXNFNPFGSPTPGVRIGTLAG 243
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      187 FSALLAELAEANRACGIPAGVFNVVTSAGAVGEBLTSPULVKRUSFTGSTBTIGROLMEQA 246
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      244 --MRPIMTELGGKDAGIVADADLDLNAKOIYVAGVYGSGORCTAIKRVLVVEEVADELA 301
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      247 KDIKKYSIELGGNAPIVFDADDLDRAKVBGALASKPRMNGQTCCVANRLVVOGGVDRFA 306
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      302 EKISENVAKLSTVGDP--DNAITYTPVIDDNNSADPIESLVYDAROKAKELNEFK--RDGR 357
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      307 EXLOOAVEELRGIDGLQDDGVTTGPLLIDEKAVAVVEBHIDAIAKKGAKVYTGKPRHALGEN 366
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      358 LITPGLFHDVHTDMKLAMEEPFCPIPIITRVKDAEEVAIALNKSDGLOS SVTRPDPOKA 417
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      367 PFQPFIIVVNPSPSAVAKETFOGLAPLRFRRKOBADVIAQANDTBEGLLAIFYARDLSRV 426
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      418 FDIANKLEVGIVHIINNTKG-RGBDNPFPLGLKSGSAGVOGIRYSIEAMTNVKSIVLDM 474
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      427 FRVGALEVGIIIGIN--TGIIISTEVAPFEGSVKASGIGRESGSKGIEUYEIKMCJGL 482
```

```

RESULT 59
ABO66116
ID ABO66116 standard; protein; 487 AA.
XX
AC ABO66116;
XX
DT 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polypeptide seqid 12633.
XX
KM Recombinant expression vector; transcription regulatory element;
KM Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
OS Klebsiella pneumoniae.
XX
PN US6610836-B1
XX
PD 26-AUG-2003.
XX
PF 27-JAN-2000; 2000US-00489039.
XX
PR 29-JAN-1999; 99US-0117747P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL, Osborne M;
XX
DR WPI; 2003-895346/82.
DR N-PSDB; ACH99667.
XX
PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX
PS Disclosure; SEQ ID NO 12633; 932pp; English.
XX
PC The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition

```

CC against *Klebsiella pneumoniae*. This is the amino acid sequence of a
CC
CC *Klebsiella pneumoniae* polypeptide of the invention
XX
SQ Sequence 487 AA;

Query Match	29.7%;	Score 711;	DB 7;	Length 487;
Best Local Similarity	34.1%;	Pred. No. 1.8e-50;		
Matches 163; Conservative	90;	Mismatches 209;	Indels 16;	Gaps 7

```

OY 6 QNYNCEKXSVN--QEIISPIDSSIGFPAWTRREVDHAMKRGKRALPAMALTYTE 63
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 QAMIDQKRDNPNGDVAIVTNPAGBQIGSVKPKGADTBREAIIEANRLPAMRLTAKE 76
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 64 PAOYLHKAADIIEERDKERIATVLAKEISKANAVASTEVVRADLIRYAEEGIRLSTSD 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 RANILRRFIDLMMENODDLARLMTLEBQKPLAAKGEISTVYASFLEWFAEBGSKRIYGDTI 136
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 124 EGGKMDASTGHKLAVIRROPVIGYLAIPVY PYNLSGSKTAPALIGNVYMFKEPTGOS 183
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 PGHAD-----KRLIVIKOPIGVPAALTIPMNPAMAIRKKGAPLMAACTWYLKPAQOTP 191
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 184 VSGVLVAKAFABAGLPGVFNPTITIRGSEIDDIYVEHEEVNPIFNTGTSPTVGORGLKLAG 243
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 FSALLALFELNARAGIPAGVENVYVTSAGAVGELTSNPLVKLSTSTGTEIGRQIMEOCA 251
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 244 --MRPIMLELGKGAGIYLAADADADINNAKQIVAGAYDYSGORCTAIRKVLVYVEADELA 301
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 252 KDIKVSILEGONNPFIVFDADDADKAVEGALSGFRNAGGTCVANNRLYODGYDIFA 311
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 302 EKISENNAKLSVGPFP--DNATVTVPIYIDNNSADFIESLNVVDARQKQAKELNEFK--RQGR 357
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 312 EKLQAAVEKRIIGGLQDQVTTGGLIDEKAAVAKVEBHADAIAGCAKVVGTGKPHALQGN 371
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 358 LLTFLGFLPHTVLLDMKLAWBEPFGFLPIIRVYKDAEBAVALINKSPFLGSSVFTFRFOKA 417
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 372 FFGQPIIVNVDSKAVAKEETFGFLAPLFRKXDEADVANOANDREFGLAATFYARDLSRV 431
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 418 FDIANKLEVGTVHINNKTG--RGPDPNFPPLGLKSGAGVQIGIRYSIEAMTNVKSIVLDM 474
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 432 FRVEALEVYGIGIIN--TGIISTVAPPGVYKAGLGBEGSKYGEDVLEIKYMOIGL 487
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT: 60
ADSD29678
ID ADS29678 standard; protein; 489 AA.
XX
AC ADS29678;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #18711.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomanan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 DR WPI; 2004-061375/06.
 XX
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 PS Claim 1, SEQ ID NO 18711, 122pp; English.
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 XX
 SQ Sequence 489 AA;
 Query Match 29.7%; Score 710.5; DB 8; Length 489;
 Best Local Similarity 35.9%; Pred. No. 2e-50;
 Matches 167; Conservative 84; Mismatches 199; Indels 15; Gaps 6;
 QY 8 YNNGEM--KSSVNGEILSPIDSSGFGVPAMTREFEVHAKRGRALPAMALTYVERA 65
 15 YIDGEVPPGDRGTLEVDPPSTELFETIPAGTTDDVAAYEAADAQEMADTPPOERA 74
 DB 66 QYIHKRAADIIEERDKETIATVTAKEISKANASTVEVRRADLRVAEGIRLSTADSG 125
 75 RIVAAASEILASHRDVLDLMTESGTTTKGAEEFAGVGTIDEAASPPTKRN-----G 129
 QY 126 GKMDASTGKHLAIVIRQPGVIVLAIPVYVNLSSGSKIAPALIGNVAMFPPTQSGVS 185
 130 DHKQSVBEKENIVAREPQGVGIIISPMFPLNLSRAVAPAAAGNSVYLKASTPTPT 189
 DB 186 -GIVAKAPAEAGLPAQVNTTIGRSEIGDYIVHEHEVNFINTSGTSPVQRIGTLAM 244
 190 GGLIAKVEAAGLPGGVFNVTGKSEIIGDRLAGPHADVISTGSTAIGVAAFTABE 249
 QY 245 RPIIM--LEIGSGDAGIVLADADLDNAAKQIVAGATVYSGORCTAIGVLYVEVADLAE 302
 250 NLALPMEIGGNVNLVTBDADLDVDSGVFSSPLHSQICISIRHVVHBEVYDEYAE 309
 DB 303 KISENVAKLSVGPDPFNATVT--PVIDNSADFIESTIVVARQGAKEINFPKRDGLTTP 361
 310 RLTERAALPTGTAHADATVVGPIIDESGRDQILDVITERTVAGALTELGCHDGLVVER 369
 QY 362 GLFDHVTLMKLAWEPPFGPIILIRVKAEEAVAIANKSDFGLQSGVFTTRDFOKADIA 421
 370 TVLSDATNMAAACHNEHFGVAVPIPFSDDEAVEIANSTSEYGLSGSVAGDTGRARDIA 429
 QY 422 NKLVEGTVINNKKTGPDNFPFLGLKSGAGVGQIRISLEMTN 466

DB 430 DRIDTGMTHINDQPNQEPHLPFGGVKNSGLG----RVNGEIVLN 470
 RESULT 61
 AD524538
 AD524538 standard; protein; 473 AA.
 AC
 AC AD524538;
 XX
 XX 02-DEC-2004 (first entry)
 DT
 DT
 DE Bacterial polypeptide #13571.
 XX
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 KW Bacteria.
 OS
 OS US2003233675-A1.
 PN
 PN 18-DEC-2003.
 PD
 PD 20-FEB-2003; 2003US-00369493.
 PF
 PF 21-FEB-2002; 2002US-0360039P.
 PR
 PR (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 PI WPI; 2004-061375/06.
 DR
 DR
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 PS Claim 1, SEQ ID NO 13571, 122pp; English.
 XX
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX

SQ Sequence 473 AA;

Query Match 29.5%; Score 705.5; DB 8; Length 473;

Best Local Similarity 34.6%; Pred. No. 4.9e-50;

Matches 167; Conservative 94; Mismatches 194; Indels 27; Gaps 9;

QY 8 YVNGEMKSSVN--QIEILSPIDSSIGFVPAMTREEDVHAMKAGREALPAMALTYVERA 65
 DB 1 YIDGQVWVSSNGDITLKYPVPTGVEYIGKFAAASRDDVDRAFAAEDAFQKMDLDSVERS 60
 QY 66 QYLHKAADIIERDKKEIATVLAKEISKAVNASVTEVVRADLRRAAEIGILSTADBG 125
 DB 61 KIIVAKELIEENRKELENIIIQENKPYKEAAEYDGLDQYAAEWARKLNGEVEG 120
 QY 126 GKMDASTGHKLAVIRROPVGIATAPYVNVNLSGSKIAPALIGNVVMFRP--PTGS 183
 DB 121 -----TTNHRKIFQYKVPYGIYVALTLPNFPAGMAREKALPALLIGNVYILKPSDTPGS 175
 QY 184 VSGVLVAKAFAPAGLPAAGVFNITIGRSGEIGDYIYHEHEVNFINTGSTPVGORIGK--L 241
 DB 176 AEWIV--KKFIOAGVPKGLANTITGSGVIGDYIYEHKKVALITMTGSTGTQRTMQKAS 233
 QY 242 AGMRIMELGCKDAGIYVLADLDNAAQIVAGAYDSGGCTAKRVLVYEEVADLA 301
 DB 224 ANMAATLMELEGKAPPMWKADINNALKTLLMAKTMNAGSCIAERLYVHEDYDPTI 293
 QY 302 EKISENVAKLSVGDPEDNATVTPVIDNSADFTESLVVDARQKAK-----ELNEBK 353
 DB 224 NKFVLSKRIIVGDP--SRADMGPLINTTLKGTBEVVEAKESGARILYGSKPLDSGPY 352
 QY 354 RDGRLLTPGLFDHVTLDMLAMEEPPGPIILPIRVDAEAVAIAKNSDFGLQSSVFTRD 413
 DB 353 KNGEFLPTIIDNVOKSKIFQEEIFAPYIGARKVSDVEEMYSLANDSKYGLASTYLFTRD 412
 QY 414 FQKADINAKLEVGVTHINNKTRGPDNRP--FLGKSGAGVQGIKRSIEMMTVKSLV 471
 DB 413 PNLIEAERIRFEGELYINMP--GPEASQGYHTGFRLLTGOAGESRGYLEYLKIKNY 469
 QY 472 LD 473
 DB 470 VD 471

RESULT 62
 ABB08382
 ID ABB08382 standard; protein; 491 AA.
 AC ABB08382;
 DT 29-AUG-2003 (revised)
 DT 18-JUN-2002 (first entry)
 DE PcuC amino acid sequence.
 KW PHBA; para-Hydrobenzoate; liquid crystal polymer; LCP;
 KM toluene monooxygenase; TMO; pcu gene; p-cresol.
 OS Pseudomonas mendocina; KR-1.
 XX WO200192539-A2.
 XX
 XX 06-DEC-2001.
 XX
 XX 22-MAY-2001; 2001WO-US016574.
 XX
 XX 01-JUN-2000; 2000US-00585174.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Ben-Bassat A, Cattermole M, Gatenby AA, Gibson KJ;
 XX Ramos-Gonzales MI, Ramos JU, Sariastani S;
 XX MPI; 2002-171436/22.
 DR

DR N-PSDB; ABA97235.

PT New nucleic acid fragments encoding bacterial toluene monooxygenase
 PT enzyme pathway, useful for isolating genes encoding proteins from the
 PT same or other microbial species, and for producing para-hydroxybenzoate.

PS Claim 4; Page 67-68; 97pp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
 CC bacterial toluene monooxygenase enzyme pathway. The nucleic acid
 CC fragments of the invention may be used to isolate genes encoding proteins
 CC from the same or other microbial species. Bacterial strains transformed
 CC with the p-cresol utilizing (pcu) genes are useful for producing para-
 CC hydroxybenzoate (PHBA) which can be used for synthesizing liquid crystal
 CC polymers (LCP). The current sequence represents the pcu amino acid
 CC sequence, which has the enzyme activity of PHBA. (Updated on 29-AUG-2003
 CC to standardise OS field)

SQ Sequence 491 AA;

Query Match 29.5%; Score 705.5; DB 5; Length 491;

Best Local Similarity 34.5%; Pred. No. 5.2e-50;

Matches 167; Conservative 94; Mismatches 196; Indels 25; Gaps 8;

QY 5 YQNTY---VNGEMK--SSVNOIEILSPIDSSIGFVPAMTREEDVHAMKAGREALPAMAA 58
 DB 8 YQNELQGLAQGWPRAGSSGRPLEVDFPNDELLRILASREDDDAVKKARDSQREWAT 67
 QY 59 LTVYRAQYLKADIIERDKKEIATVLAKEISKAVNASVTEVVRADLRRAAEIGRL 118
 DB 68 TAPERARVLEAVKIFDERREIIDWIRRESG-----TRIKQIEMGARAATL 118
 QY 119 STSADG---GKMDAST--GHKLAVIRROPVGIATAPYVNVNLSGSKIAPALIGNVV 174
 DB 119 ESASLPNKHGRIITASNISGKESRYRAPLVGITYISWNPFLHTANSLAPALAGNAV 178
 QY 175 MEKPTQGSVS--GLVAKAFAPAGLPAAGVFNITIGRSGEIGDYIYHEHEVNFINTGSTP 233
 DB 179 VYKPSADPTPIIGULLARIFEDAGLPAGVLVYVVGSGAEIDAFVHEHPVPLISPTGTD 238
 QY 234 VGORIGKLAG---MRPMLLEGCKDAGIYVLADLDNAAQIVAGAYDSGGCTAKR 289
 DB 239 VGRNIGRIASGGEHLKHALIEIGNSPFVLADLADVEQAVNAAVGKFLHGOQICMAINR 298
 QY 290 VLVYEEVADLEAKISENVAKLSVGDPEDNATVTP--PVIDNSADFTESLVVDARQKAKE 348
 DB 299 IIVEQPLLEDFTRRKVERVKALPYGDESKPGIVGVPTNARQAGLKKEKITPAKAGATL 358
 QY 349 LNEFRDGRLLTPGLFDHVTLDMLAMEEPPGPIILPIRVDAEAVAIAKNSDFGLQSS 408
 DB 359 ILGSRPQGNVMPHFVFGVVTADMEIAREIFGLVIGISARDADHALLANSSEYGLSSA 418
 QY 409 VFTRDFOKAFDIANKLEVGVTHINNKTRGPDNRPFLGLKSGAGVQGIKRSIEMMTVK 468
 DB 419 VFTASLBERGVQFARRIHAGMTHVNDIPVNDPNAPFGGKXSGLGRFNGDWAIBERTTDH 478
 QY 469 SIYL 472
 DB 479 WITL 482

RESULT 63

ABU92062

ID ABU92062 standard; protein; 491 AA.

AC ABU92062;

DT 15-JUL-2003 (first entry)

DE Pseudomonas mendocina p-cresol utilizing protein C (pcuC).

KW TmoST; tmoS; toluene monooxygenase; para-hydroxybenzoate; PHBA;

KM toluene degradation pathway; TMO; p-cresol utilizing; pcu; operon.

```

XX OS Pseudomonas mendocina.
XX PN US002151003-A1.
XX PD 17-OCT-2002.
XX PF 28-NOV-2001; 2001US-0097664.
XX PR 01-JUN-2000; 2000US-00585174.
XX PA (BENB/) BEN-BASSAT A.
XX PA (CATT/) CATTEMOLE M.
XX PA (GATE/) GATENBY A. A.
XX PA (GIBS/) GIBSON K. J.
XX PA (RAMO/) RAMOS-GONZALEZ M. I.
XX PA (SARI/) SARIASLANI S.
XX PI Ben-Bassat A, Cattemole M, Gatenby AA, Gibson KJ;
XX PI Ramos-Gonzalez MI, Ramos JL, Sariaslani S;
XX DR MPI; 2003-428753/40.
XX DR N-PSDB; ACA92456.
XX PT Novel nucleic acid sequence encoding Tmost polypeptides, useful for
XX PT obtaining a nucleic acid fragment encoding Tmost polypeptides.
XX PS Example 1; Page 27-29; 65pp; English.
XX CC The invention describes an isolated nucleic acid fragment (I) of
XX CC Pseudomonas mendocina KR-1, comprising a 4821 nucleotide sequence and
XX CC encoding a Tmost polypeptide (referring to toluene monooxygenase
XX CC polypeptides, Tmost comprising 973 amino acids (a.a.s) having at least 83%
XX CC identity and Tmost comprising 220 a.a.s having at least 85% identity based
XX CC on the Smith-Waterman method of alignment with a sequence). (I) is useful
XX CC for obtaining a nucleic acid fragment encoding a Tmost polypeptide, by
XX CC probing a genomic library with (I), selecting for a DNA clone that
XX CC hybridizes with (I), and sequencing the genomic fragment that comprises
XX CC the identified clone, where the sequenced genomic fragment encodes Tmost
XX CC polypeptide. The polypeptide encoded by (I) is useful as a tool to
XX CC mediate expression of the catabolic two genes and para-hydroxybenzoate
XX CC (PHBA) production in any organism that does not possess (I). This is the
XX CC amino acid sequence of a protein encoded by the Pseudomonas mendocina p-
XX CC cresol utilizing (pcu) gene operon the proteins encoded by which are
XX CC involved in the toluene degradation pathway (TMO)
XX SQ Sequence 491 AA;
Query Match 29.5%; Score 705.5; DB 6; Length 491;
Best Local Similarity 34.5%; Pred. No. 5.2e-50;
Matches 167; Conservative 94; Mismatches 198; Indels 25; Gaps 8;
QY 5 YONY-----NGEKK--SSVNOIEILSPIDSSLGFPNMTREVNDAMKAGRALPAMA 58
DB 8 YONLEIPIAGQGRRSSGRPLEVDPYDELRLIALSRDLDAAYRKARDSQREWAT 67
QY 59 LTVYERAOYLHKAADIIEEDKEIATVLAKEISKAYNAVTEVFRADILRYAEEGRL 118
DB 68 TAPAEARVULEVVKIFDRRERIIMTIRESS-----TRKAQIEMWAAALITL 118
QY 119 STSADG---GKMDAST-GHKLAIVRRQVGLAIPYNYPNLSGSKIAPALIGNAV 174
DB 119 ESSSLERNYHGRILIANISGKESRVRAPGVISPMFPIHLTARSIALPALALGNV 178
QY 175 MFAPPTQGSVS-GLVYAKAPAEAGLPAGVFNITTGSGSEIGDVIYHEEVPNFNTGSP 233
DB 179 VAKPASDPTPTGGLLARIFFEEGLPAGLVSVVGSAEIGDFVHPVPALISFTGSIQ 238
QY 234 VGGRIKLAG---MPEIMLELGGKDAIVLADADDLNNAKQIVAGAYDYSGQRCTAIR 289
DB 239 VGNIGRIASGGEHLKHALVELGSGNSPPVTLADADVQAVNAVVKFLHQGGICMAINR 298

```

```

QY 290 VLVEEVADELAEKISENVAKLSVGPPDNATVYT--PVIDNSAGDFIESLWVDAPOKAKE 348
DB 299 LIVEQPLEDFTRRVERVKALPYGDPKRGTVGEPVVARQLAGUKEXIARAKAGATL 358
QY 349 INEFKRDGRLTPGLFDHVTLDMKLAMEEPFGIILPIIRKDAEEAVALANKSDPGLQSS 408
DB 359 ILGGEQGNVMPPHVGNVATAMEIAREBEIPPLVIGISGARAHEHLEANSSEYLSGA 418
QY 409 VTRDFQKAFDIANKLEVGTHIINKTGRGPDNFPPLGLKSGAGVQGIKRISEAMTNVK 468
DB 419 VFTASLERGVQFARRIHAGMTHVNDIPVNDENAPAFGGERKNSGLGRFNGDMAIEEFTDH 478
QY 469 SIYL 472
DB 479 WITL 482
RESULT 64
ABU40285
ID ABU40285 standard; protein; 480 AA.
XX AC ABU40285;
XX AC ABU40285;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #25812.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Pseudomonas putida.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR MPI; 2003-029926/02.
XX DR N-PSDB; ACA44155.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 68209; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 623 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an

```

organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 480 AA;

Query Match 29.4%; Score 704; DB 6; Length 480;
Best Local Similarity 34.0%; Pred. No. 6,7e-50;
Matches 162; Conservative 91; Mismatches 209; Indels 14; Gaps 7;

6 QNYVNGEMKSSVNV--QIELSLPIDDSLGFPVPMTEEDVHMKAREALPMAALTVE 63
12 QYVINGEMWDADNGQTITVNPATGVEIVTVPKMGTAETRRALAEADKALPMWRLTAKE 71
64 RAQYLHKADIIERDEKEIATVLAKEISKAVNASVTEVVRTADLIRYAEBSGIRLSTSD 123
72 RBAKLIRWFELMIENDDLARLMTTEOGKPLAEKEIYVAAFSFIMWPEAEKRIYGDIT 131
124 EGGKMDASTGHHKLAIVIRQPVGIVLAIDYNTPVNLSSKIAIPALIGVNVVFKPPTGGS 183
132 PGHQP-----RLIYIKQPIGVTAITPMNFPAMITRKAGPALAAGCTMVLKPAQTTP 186
184 VSGVLAKAFAPAGVFNITTGSGSRIQDIYVHEBVNFINTGSPVQGRICK--L 241
187 YALALVELAHRRGIPAGVLSVVTSGAGEVGGELTGNLSVRKLSFGSTIEGQLMECA 246
242 AGWRPIMLEBGKNDAGIVLADLDLNAAQIVAGADYSGORCTAKRLVVEEVADELA 301
247 KQIKKVSLELGGNAPRTFDDADLDKAVEGAIISKRNNGQTCVCAKRIYVQDGYDARA 306
302 EKISENVAKLSVGDPEPDNAVT--PVTDNSADFTESLVDVARQKAKEINERK--DGRLL 359
307 EKLAAVAVALKIKNGIEEGTITGPLLDGKAVAKVQHIDAVSKKAVISGGKLTIEGNFF 366
360 TFGLPDHYVLDMKLMEEBPGPLPIIRYKDAEVAIVANKSPFGQSSEVTRDPQKAD 419
367 EPTILVDVEKTAIVAKEBETFGPLAPLFRFDEAEVITAMSNDEFGIASYFYVADMSRVFR 426
420 IANKLEVGTVAHINNKTR--GPDNFPPLGKSGAGVQIGRYSIEANTNYSIVLDM 474
427 VAEALEYGVNGIN--TGLISNEVAPRGGIKASGLGREGSKYGEIDYLEIKYLCISV 480

RESULT 65
ADS28136
ID ADS28136 standard; protein; 468 AA.
XX
AC ADS28136;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #17169.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX

Bacteria.
US2003233675-A1.
18-DEC-2003.
20-FEB-2003; 2003US-00369493.
21-FEB-2002; 2002US-0360039P.
(CAOY/) CAO Y.
(HINK/) HINKLE S J.
(SLAT/) SLATER S C.
(CHEN/) CHEN X.
(GOLD/) GOLDMAN B S.
Cao Y, Hinkle SJ, Slater SC, Chen X, Goldman BS;
WPI; 2004-061375/06.
New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.
Claim 1; SEQ ID NO 17169; 122pp; English.
The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polypeptide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 468 AA;

Query Match 29.1%; Score 697.5; DB 8; Length 468;
Best Local Similarity 33.4%; Pred. No. 2.3e-49;
Matches 158; Conservative 93; Mismatches 209; Indels 13; Gaps 5;

8 YVNGEMKSSVNVQIELSLPIDDSLGFPVPMTEEDVHMKAREALPMAALTVEYROY 67
4 YVNGEMKSSGKTLVDVNPATGVEIVTVPAGKKEALVSAAYEAPFQWVMEAKRI-----YDDT 63
68 LHKADIIERDEKEIATVLAKEISKAVNASVTEVVRTADLIRYAEBSGIRLSTSDBSGK 127
64 LMRWFQILDQGDDELGEIMTEKGKPLAEALGEVQYANFSIQWVMEAKRI-----YDDT 118
128 MDASTGHHKLAIVIRQPVGIVLAIDYNTPVNLSSKIAIPALIGVNVVFKPPTGSGSL 187
119 IPASAINRKRIIVQKQPVGVIAAITPMNFPAMITRKVAPALAAAGCTAIVKPAEQTPPAL 178
168 VLAKAFAPAGVFNITTGSGSRIQDIYVHEBVNFINTGSPVQGRICK--GLAAGR 245
179 KLAQLAEBAGIPAGVLSVVTGNAODIGAMLEDSRKRKITFGSTIEVGLKLRGAQIVTK 238

Oy		246 P.M.L.E.G.C.D.A.G.I.V.I.A.D.A.L.D.N.A.A.K.O.Y.A.G.A.N.Y.S.S.O.R.C.T.A.I.R.V.L.V.E.E.B.A.B.E.L.A.E.K.I.S 305
Dd		229 K.I.S.L.E.I.G.G.A.P.P.I.I.M.D.D.A.N.L.E.E.V.D.V.I.A.S.K.F.F.N.N.A.Q.O.T.C.C.A.N.R.I.Y.V.A.R.E.I.A.F.A.F.E.K.F.A 298
Oy		306 E.N.V.A.K.I.S.V.D.D.P.-D.N.A.T.V.P.V.I.D.D.N.S.A.D.F.I.E.S.I.V.V.A.R.Q.K.A.E.L.N.E.F.F.K.D.G.L.T.P.G.L.F 364
Dd		229 A.K.V.N.E.L.K.V.N.G.L.E.E.G.T.T.I.G.P.L.D.K.A.V.E.K.E.A.H.I.H.A.L.K.K.G.O.V.Y.T.G.R.W.T.N.H.F.F.E.P.T.I.I 358
Oy		365 D.H.V.T.L.D.M.K.L.A.W.E.E.P.F.G.P.I.L.P.I.R.V.K.D.A.E.E.A.V.A.L.A.N.K.S.D.P.G.L.O.S.S.V.F.T.R.D.F.O.K.A.F.D.I.A.N.K.L 424
Dd		359 T.G.A.T.E.M.L.C.M.N.E.E.T.F.C.P.L.A.P.V.A.T.F.D.E.E.V.I.E.R.A.N.H.T.P.G.L.A.A.V.F.F.E.N.I.G.R.A.I.R.I.S.E.K.L 418
Oy		425 E.V.G.T.Y.H.I.N.N.K.T.G.R.P.D.-N.F.P.F.L.G.K.G.S.G.A.N.G.O.C.I.Y.S.I.E.A.M.N.V.K.S.I.V.I.D.M.K 475
Dd		419 E.X.G.I.V.G.V.N.D.G.M.--P.S.V.A.Q.A.P.F.G.O.W.K.E.S.G.I.G.R.E.G.K.Y.G.I.E.E.Y.L.E.V.Y.V.S.V.A.F.K 468
		RESULT 66
		ADSS29963
ID	ADSS29963	standard; protein; 473 AA.
XX	AC	
XX	ADSS29963;	
XX	XX	
DT	02-DEC-2004	(first entry)
XX	DE	
XX	XX	Bacterial polypeptide #18996.
KW	KW	Recombinant DNA construct; transformed plant; improved plant property;
KW	KW	cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
KW	KM	pathogen tolerance; pest tolerance; plant disease resistance;
KW	KM	cell cycle pathway modification; plant growth regulator;
KW	KV	homologous recombination; seed oil yield; protein yield; carbohydrate;
KW	KV	nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX	XX	bacterial polypeptide.
OS	XX	Bacteria.
XX	PN	
PN	US2003233675-A1.	
PD	18-DEC-2003..	
XX	XX	
PF	20-FEB-2003; 2003US-00369493.	
XX	PR	
PR	21-FEB-2002; 2002US-0360039P.	
XX	CAO Y.	
PA	(HINK)/ HINKLE G J.	
PA	(SLAT)/ SLATER S C.	
PA	(CHEN)/ CHEN X.	
PA	(GOLD)/ GOLDMAN B S.	
XX	XX	
PI	Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;	
WR	WI, 2004-061375/06.	
PT	XX	
PT	XX	New recombinant DNA construct comprising a promoter positioned to provide
PT	XX	for expression of a polynucleotide encoding a polypeptide from a
PT	XX	microbial source, useful for producing plants with improved properties.
PS	Claim 1; SEQ ID NO 18996; 122bp; English.	
XX	XX	
CC	XX	The invention relates to a recombinant DNA construct comprising a
CC	CC	promoter functional in a plant cell, where the promoter is positioned
CC	CC	to provide for expression of a polynucleotide encoding a polypeptide from a
CC	CC	microbial source. The invention also relates to a transformed plant
CC	CC	comprising the recombinant DNA construct and a method of producing a
CC	CC	transformed plant having an improved property. The plant is a crop plant
CC	CC	such as maize or soybean. The method of producing a transformed plant
CC	CC	having an improved property comprises transforming a plant with the
CC	CC	recombinant DNA construct and growing the transformed plant, where the
CC	CC	polynucleotide or polypeptide is useful for improving plant properties.
CC	CC	The recombinant DNA construct is useful for producing plants with
CC	CC	improved plant properties, e.g. improved cold, heat or drought tolerance,

	CC	tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
	CC	increased resistance to plant disease, better growth rate by modification
	CC	of the cell cycle pathway with plant growth regulators, increased rate of
	CC	homologous recombination, modified seed oil or protein yield and/or
	CC	content, improved yield by modification of carbohydrate, nitrogen or
	CC	phosphorus use and/or uptake, by modification of photosynthesis or by
	CC	providing improved plant growth and development under at least one stress
	CC	condition, improved lignin production or improved galactomanan
	CC	production. This sequence represents a bacterial polypeptide used in the
	CC	scope of the invention. Note: The sequence data for this patent did not
	CC	form part of the printed specification but was obtained in electronic
	CC	format from USPTO at seqdata.uspto.gov/sequence.html.
XX	Sequence 473 AA:	
SQ	Query Match	29.0%; Score 694; DB 8; Length 473;
	Best Local Similarity	32.6%; Pred. No. 4, 5e-49;
	Matches 157; Conservative 102; Mismatches 198; Indels 24; Gaps 8;	
OY	8 YVNGEWSVNOJIEILS--PIDDSLT-GFVPAMTREBEVDHAKAGREALPAMAALTVYER 64	
Dd	1 YINGOMLNAATETTLNHSNPADKSEIVATFPSSGADDVDVAVAARQAYSWMKVPAAPAR 60	
OY	65 AQYLHKRADIIIERDKSEIATVLAKESIKAVNASYTEVEVRPADLRIRAAEGCIRLSADE 124	
Dd	61 AEYIFRGEELLOHKEELADLIISRMGKPLTEARGDVQEBGVDAFYSGAGEGRLLF----- 115	
OY	125 GGRKDASTGHKLAVIRGPVGIYLAIPVNVYPNLSSSKTAPRLIGNVVMFPPIQGVS 184	
Dd	116 GQTTPSPMPNKFAATMMPIGVCALLTPMNPVPAVIPCKAMPALVCGNTVILRPADETPA 175	
OY	185 SGLVLAFAFAEAGLPAGVFNTITGSGEIDGYVEHEEVAFINFTGSTPVGORIGKLGM 244	
Dd	176 CATLIIIFPAAGLPPEVINLVHGVSVEGKALVENHNIDLVEFTSSAAGAAYGETCGR 235	
OY	245 --RPIMLELGKQDAGIVLADADLDNNAKQIVAGAYDYSGGRCRAIKRVLVVEEVADELA 302	
Dd	236 THKVKCLEMGCKNAQOVMEADALELADGALMGAFGTTCORCTATSRLILHRDIKERFTT 295	
OY	303 KISBNVAKLSVGDFPDNAT-VTPRIIDNSADFIESLVVDNRQGAKEI--NEFRDQRL- 358	
Dd	296 MLRERTSQRLRGAGTEBETDIGPINNRQLQRHYENNNIAREBAKLLIGGETATEBOLQ 355	
OY	359 ----LTPEGLFDHYVTLDDKLAMEEPFGELIPIIRKYDAEEVAIALANKSDPFLQSSVPRDF 414	
Dd	356 QGYFGQTIIDNTYTPQKVARBELFGVVALLIVSTFEELAIILNDPKYGLSSSVYTRDI 415	
OY	415 QKAEDIANKLVEGTVAHINNKTGSGPPDPFPLGLKGSGAGVOGIRYSIEAMTNVKSIVLDM 474	
Dd	416 NRARVANRDIAGITVINGPITIGAETHVLPFGVGKQKONGHR-----EAQTTALDVETEW 469	
OY	475 K 475	
Dd	470 K 470	
RESULT 67		
ID	ADN26088	
XX	ADN26088 standard; protein; 477 AA.	
AC	ADN26088;	
XX		
DT	02-DEC-2004 (first entry)	
XX	Bacterial polypeptide #8741.	
KW	Recombinant DNA construct; transformed plant; improved plant property;	
KW	cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;	
KW	pathogen tolerance; pest tolerance; plant disease resistance;	
KW	cell cycle pathway modification; plant growth regulator;	
KW	homologous recombination; seed oil yield; protein yield; carbohydrate;	
KW	nitrogen; phosphorus; photosynthesis; lignin; galactomannan;	
KW	bacterial polypeptide.	

XX OS Bacteria.
 XX PN US200323675-A1.
 XX PD 18-DEC-2003.
 XX PF 20-FEB-2003; 2003US-00369493.
 XX PR 21-FEB-2002; 2002US-0360039P.
 XX PA (CAOY/) CAO Y.
 XX PA (HINK/) HINKLE G J.
 XX PA (SLAT/) SLATER S C.
 XX PA (CHEN/) CHEN X.
 XX PA (GOLD/) GOLDMAN B S.
 XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX DR WPI; 2004-061375/06.
 XX PT New recombinant DNA construct comprising a promoter positioned to provide
 XX PT for expression of a polynucleotide encoding a polypeptide from a
 XX PT microbial source, useful for producing plants with improved properties.
 XX PS Claim 1; SEQ ID NO 8741; 122pp; English.
 XX CC The invention relates to a recombinant DNA construct comprising a
 XX CC promoter functional in a plant cell, where the promoter is positioned to
 XX CC provide for expression of a polynucleotide encoding a polypeptide from a
 XX CC microbial source. The invention also relates to a transformed plant
 XX CC comprising the recombinant DNA construct and a method of producing a
 XX CC transformed plant having an improved property. The plant is a crop plant
 XX CC such as maize or soybean. The method of producing a transformed plant
 XX CC having an improved property comprises transforming a plant with the
 XX CC recombinant DNA construct and growing the transformed plant, where the
 XX CC polynucleotide or polypeptide is useful for improving plant properties.
 XX CC The recombinant DNA construct is useful for producing plants with
 XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 XX CC increased resistance to plant disease, better growth rate by modification
 XX CC of the cell cycle pathway with plant growth regulators, increased rate of
 XX CC homologous recombination, modified seed oil or protein yield and/or
 XX CC content, improved yield by modification of carbohydrate, nitrogen or
 XX CC phosphorus use and/or uptake, by modification of photosynthesis or by
 XX CC providing improved plant growth and development under at least one stress
 XX CC condition, improved lignin production or improved galactomannan
 XX CC production. This sequence represents a bacterial polypeptide used in the
 XX CC scope of the invention. Note: The sequence data for this patent did not
 XX CC form part of the printed specification but was obtained in electronic
 XX CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX SO Sequence 477 AA;

Query Match 29.0%; Score 694; DB 8; Length 477;
 Best Local Similarity 35.5%; Pred. No. 4,5e-49;
 Matches 173; Conservative 83; Mismatches 197; Indels 34; Gaps 10;

QY 5 YGNVNGEKKSSVNOEILSPIDSSLGVPNATREVP-HAMKAREALPA----- 55
 DB 3 YRNYIDGECWSASG---RTVDNVN---PADTRIVSRHSAASDRAAAAAA 54
 QY 56 --MAALTVERAQYLKKAADIERDKKEIATVLAKISKAYNVAVEVETADLIYAAE 113
 DB 55 DGMKPTPKRKAKILNDAAAHLEANDTATLTREGRALNARBEVRSQOTLAFYAV 114
 QY 114 EGIRLSTADBEGRKMDASTGHLKAVIRROPVGIVLAIPYNPVNLGSKIADILGAV 173
 DB 115 EGQTFGSGESYPNDPD-----MLVYSLREPLAGVTVISPMNPFVSIARKIAPALITGNT 169
 QY 174 WNEKPTQGSVSLVAKAFAEAGLPAQVFNITTGSGSEIGDYIYHEEFPNTNFGSGTP 233
 DB 170 VVFKSSDAPLDSGYRLAEAFVRAGIRKGVNLFLTGSAAAEVGPVIVSRREAVSFTGSGTS 229

QY 234 VGORIGKLAKMRP-IMLEIGGKDAGIVLADDLNPAKQIVAGAYVSGORCTAIKRVLY 292
 DB 230 AGEQIHKSVPMTTRTQMLGGKNPILVIMEDADLDRAVDLAVKGGSLSGQCTGSRLLV 289
 QY 293 VEEVADLEAKTISEVNAALSVGDDPD-NATTPVIDNDSAPFISBLVVDARQ-----GA 346
 DB 290 MEDVQAFTKELVAKVAKALKIGSGMTAGMDLGPATRKQELTVLSYIAIGQEAATLGG 349
 QY 347 KEINFEK-RDGRLLRPGFHDVYTLDMKLAWEPPGPIITIRYKDAEAAVALNANKSDRGL 405
 DB 350 EQLSDGDRAHGYVAPVFTDTQSMRIAREBIFEPVALIIEVDSYADALNADTEGL 409
 QY 406 QSSVPTRDFOKAFDIANKLEVGTVHINNKTRGPDNPFLGKSGAGV--QGIRYSIEA 463
 DB 410 SAAIATRPVYVHMDPTRDIESGTIVKINTTGNLVNAPFGGLKRSSTSPFESGRAGLEF 469
 QY 464 MTNFKSI 470
 DB 470 YTOIKTV 476

RESULT 68
 ABU41623
 ID ABU41623 standard; protein; 480 AA.
 AC ABU41623;
 XX DT 19-JUN-2003 (first entry)
 XX DE Protein encoded by Prokaryotic essential gene #27150.
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX OS Pseudomonas syringae.
 XX PN WO20027183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002WO-US009107.
 XX PR 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948993.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.
 XX PR 06-MAR-2002; 2002US-0362699P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX DR WPI; 2003-029926/02.
 XX DR N-PSDB; ACh45493.
 XX PT New antisense nucleic acids, useful for identifying proteins or screening
 XX PT for homologous nucleic acids required for cellular proliferation to
 XX PT isolate candidate molecules for rational drug discovery programs.
 XX PS Claim 25; SEQ ID NO 69547; 176pp; English.
 XX CC The invention relates to an isolated nucleic acid comprising any one of
 XX CC the 623 antisense sequences given in the specification where expression
 XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
 XX CC encoding a polypeptide whose expression is inhibited by the antisense
 XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 XX CC polypeptide or its fragment whose expression is inhibited by the
 XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
 XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 XX CC proliferation or the activity of a gene in an operon required for
 XX CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 480 AA;

Query Match 29.0%; Score 693; DB 6; Length 480;

Best Local Similarity 33.8%; Pred. No. 5.5e-49;

Matches 161; Conservative 89; Mismatches 212; Indels 14; Gaps 7;

QY 6 QNYVNGEM--KSSVNOIEILSPIDSSLGFPAMTBEVDHAKGREALPMAALTYVE 63
 DB 12 QATNGQMDADGQGISIKVNNPATNELIGTVPMGAAETRRALIEADKLPARALITAKE 71
 QY 64 RAQYLHKADIIERDKSEIATVLAKEISKAYNASVTEVVRTADLIRYAEGRISTSD 123
 DB 72 RGNKLRWFELMTENQDDGLMTLEQKFLAEKGEITYAASFIEFPAEAKRYVGDVI 131
 QY 124 EGGKMASTGHKLAIVRROPVGTVALAPYNPVNLGSGTIAALGSGNVMPKPTQGS 183
 DB 132 PGHQP---KRLIVAKOPIGVTAATPNNFPAMTTRAGPLAAGCTWVLPKPSQTP 186
 QY 184 VSGQLVLAKEAFEAAGLPAGVFNITGRSGSEIGDIYVEHEEVNFINFGSTPVGORIGK--L 241
 DB 187 FSLALAEALAEKRGIPAGVSVTGSAGDIGSELTGNPIRKISFTGSGTIGQMAECA 246
 QY 242 AGNRPMLEVGSKDAGIVLADADLNNAKQIVAGAYDSGQRTAKRVLVEEVADELA 301
 DB 247 KDIKTVSELGNAPIVFDPDADLDKRAVEGAMTSKRYNNGQTCVCAIRIVYQGVDAFA 306
 QY 302 EKISENVAKISVQDPF--DNATVTPVVDNADPIESLVNARQKAKELNEFKR--DGRLL 359
 DB 307 EKKVAVGKLIKNGLEDGITTPBLIDEKAVAVKEHIDAVSKGATVLVGNSLRSFP 366
 QY 360 TPGQLFHVTLDMKLAMEEPFGPILPITRVDAEAVANAKSPDGLQSGSVTRDFOKAPD 419
 DB 367 EPIIVNVSDAIVARETFRGLAPLFRPDEBAEALITANDTFRGLASYVAQNMGRVFR 426
 QY 420 IANKLEVTGHINNKTRGPDNF--PFLGLKSGAGVQGIIRYSIEAMTNVKSIVLDM 474
 DB 427 VAAALEGVWGIN--TGLISNELAPFGIGIKSGIGREGSKYGIEDYLEIKYLCISV 480

RESULT 69

AB042152 standard; protein; 480 AA.

ID AB042152

AC AB042152;

XX 19-JUN-2003 (first entry)

DE Protein encoded by prokaryotic essential gene #27679.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Pseudomonas syringae.

XX MO200277183-A2.
 PN 03-OCT-2002.
 PD 21-MAR-2002; 2002WO-US009107.
 PF 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI MPI: 2003-029926/02.
 DR N-PSDB; ACA46022.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 70076; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SO Sequence 480 AA;

Query Match 29.0%; Score 693; DB 6; Length 480;

Best Local Similarity 33.8%; Pred. No. 5.5e-49;

Matches 161; Conservative 89; Mismatches 212; Indels 14; Gaps 7;

QY 6 QNYVNGEM--KSSVNOIEILSPIDSSLGFPAMTBEVDHAKGREALPMAALTYVE 63
 DB 12 QATNGQMDADGQGISIKVNNPATNELIGTVPMGAAETRRALIEADKLPARALITAKE 71
 QY 64 RAQYLHKADIIERDKSEIATVLAKEISKAYNASVTEVVRTADLIRYAEGRISTSD 123
 DB 72 RGNKLRWFELMTENQDDGLMTLEQKFLAEKGEITYAASFIEFPAEAKRYVGDVI 131

QY 124 EGGKMDASTGHTKLAIVIRROPVGIIVLAIAPYVFNLSGSKIAIPALIGNVVMEKPPYGS 183
DB 132 PCHQPD-----KRLIVLKQPIGVTAITPMNFPAAITRKAGALAGCTMVLKPSQTP 186
QY 184 VSGVLAKAFABAGLPAGVFNITGSGSEIGDIYVEHEVNFINTGSPVGORIGK--L 241
DB 187 PSALALAEABERAGIPAGVSVTGSAGDISGELTNPYRKLSFTGSGTEIGRQMAEQA 246
QY 242 AGMRPIMLELGGKDAIGVLADADLDNAAKOIVAGAYDVSQGRCTAIKRVLVVEEVADELA 301
DB 247 KOIKKVSLELGGNAPFIVDDADLDRAVAGAMISKYRNNGQTCVCMRIYVDGVDAFA 306
QY 302 EKISENVAKLSVGDPF--DNATVTPVIDNSADFIESLVVDAROKAGKEINEFRK--DGRLL 359
DB 307 EKLKVAIVGKLIKINGLEDGITTGRLIDEKAVAKVKEHIAVASKATVLTGNSLEGSPF 366
QY 360 TGLFPHVTLDMKLAMEEPPGPIPLPIIRKYDAEAVAIANKSDPGLOSSVFTTRDFOKAD 419
DB 367 EPTLVNVKSDAAVAAEEFTFGPLAPLFRFDEAEALANDTEFGLASFYAQNMSRVR 426
QY 420 IANKLEVGTVHINNKTKGRGPDNF--PFLGKSGSAGVQGIIRYSIEAMTNVKSIVLDM 474
DB 427 VAEALAEYGVGIN--TGLISNELAPPGIKSGLSGREGSKYGIEDYLEIKYLCISV 480

RESULT 70
ABU47293
ID ABU47293 standard; protein; 482 AA.
AC ABU47293;
XX 19-JUN-2003 (first entry)
XX Protein encoded by prokaryotic essential gene #32820.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Salmomella typhimurium.
XX WO20027183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zyekind WJ,
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX MPI: 2003-029926/02.
XX N-PSDB; ACAS1163.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 75217; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 482 AA;

Query Match 28.9%; Score 692; DB 6; Length 482;
Best Local Similarity 33.3%; Pred. No. 6.8e-49;
Matches 159; Conservative 95; Mismatches 208; Indels 16; Gaps 7;

QY 6 QNYVNGEKKSS--VNOIEILSPIDSSLSGFYPMATREBEVDHAKMGREALPMAALTYVE 63
DB 12 QAFIDGDRDARGGVIVPSNPANGKPLGNPKYGAEBETRDIAINANALPAMRLTAKE 71
QY 64 PAQYVHKADITIERDKIEIATVLAKESKAVNASVTEVVRADLIRVAEGGIRLSTAD 123
DB 72 RANILRRWFNLMMEHQDDLRLMTLEQCKPLAEAKGEISYAASTFEMFAEEGKRIVDPI 131
QY 124 EGGKMDASTGHTKLAIVIRROPVGIIVLAIAPYVFNLSGSKIAIPALIGNVVMEKPPYGS 183
DB 132 PCHQPD-----KRLIVLKQPIGVTAITPMNFPAAITRKAGALAGCTMVLKPSQTP 186
QY 184 VSGVLAKAFABAGLPAGVFNITGSGSEIGDIYVEHEVNFINTGSPVGORIGKLAG 243
DB 187 PSALALAEABERAGIPAGVFNITGSGSEIGDIYVEHEVNFINTGSPVGORIGKLAG 246
QY 244 --MRPIMLELGGKDAIGVLADADLDNAAKOIVAGAYDVSQGRCTAIKRVLVVEEVADELA 301
DB 247 KOIKKVSLELGGNAPFIVDDADLDRAVAGAMISKYRNNGQTCVCMRIYVDGVDAFA 306
QY 302 EKISENVAKLSVGDPF--DNATVTPVIDNSADFIESLVVDAROKAGKEINEFRK--DGRLL 359
DB 307 EKLKVAIVGKLIKINGLEDGITTGRLIDEKAVAKVKEHIAVASKATVLTGNSLEGSPF 366
QY 360 TGLFPHVTLDMKLAMEEPPGPIPLPIIRKYDAEAVAIANKSDPGLOSSVFTTRDFOKA 417
DB 367 FPGPITLADVPDMAVAEEFTFGPLAPLFRFDEAEALANDTEFGLASFYAQNMSRVR 426
QY 418 FDIANKLEVGTVHINNKTKGRGPDNF--PFLGKSGSAGVQGIIRYSIEAMTNVKSIVLDM 474
DB 427 FVAGALAEYGVGIN--TGLISNEVAPPGIKSGLSGREGSKYGIEDYLEIKYLCISV 482

RESULT 71

ADSS24924
ID ADSS24924 standard; protein; 478 AA.
AC ADSS24924;
XX 02-DEC-2004 (first entry)
XX Bacterial polypeptide #13957.
XX

KM Recombinant DNA construct; transformed plant; improved plant property;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KM bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 13957; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 CC
 XX
 SQ Sequence 478 AA;
 XX
 Query Match 28.2%; Score 691; DB 8; Length 478;
 Best Local Similarity 33.5%; Pred. No. 8.1e-49;
 Matches 159; Conservative 92; Mismatches 209; Indels 14; Gaps 7;
 XX
 QY 6 QNTVNSEKSSV--QETISPTDDSSLGVPVPMTEBVDHAKAGREALPMAALTYYB 63
 DB 11 QADIDAWDADGQITKVNNPTGILGTPGGAETRRALBADKALPARALTAAR 70
 QY 64 RAQYLKHAADIRDEKEITVLAKEISKAYNASVTEVVTADLIRYAAEGRTLSAD 123
 DB 71 RACKLRWPELMTENDDLARMTLLOGKRLAKAGKEIVYAAAFIEMFAEAKRIYGDVI 130

QY 124 EGGKMDASTGHLAVIRROPGVIGLAIAPVYVNLSSGKIAPALIGNVWMEKPTQGS 183
 DB 131 PCHQDP-----KRLVITKOPIGVTAITPMNPPMAITRKAGPALAGCTMWLKPASQTP 185
 QY 184 VEGLVLAFAFAAGAPAGVFNITGSGSEIGDIYVEHEVNEINFTSGTPVGRIQK--L 241
 DB 186 FSAFALAEIAGPAGAPKGVFSVSGSAGDIGSELINSPYVRKLSFTGSGIEIRQLMSSEA 245
 QY 242 AGMRPIMELEGGKADGIVLADADLDNAKQIVAGAYDSGQRCTAIRLVVVEVADELA 301
 DB 246 KDIKKVSIELGGAAPFIVFDDADLDKAVGALISKRNNGQTCVCANRLYIDSVYDARA 305
 QY 302 EKISENVAKLSVGDPEPDNATVT--PVYDDNSAPFIESLVVDAPQKAKEINERK-RGRLL 359
 DB 306 EKLKVAVAALKIINGLEBGTITGPILDERAKAVKVEHIDAVSKATVLAGKPEMGNEF 365
 QY 360 TPGLPDHTVLDMLAMEBEPGPILPIIRKDAEBAVALNKSDFGLQSSVFTFDFOKAD 419
 DB 366 EPTILTIVPKDAVAAKEETFGPLAPLFRPKEDADVIMNSNDTREFGLASIFYARDLGRVIR 425
 QY 420 IANKLEVGTVHINNKTR-GPDNPPPLGLKSGAGVGIRYSIEAMTVKSYIVL 472
 DB 426 VAEALEGMVGVN--TGLISNEVAPPGITKASGLREBSKYGIEDYLEIKYICL 477
 XX
 RESULT 72
 ABU23300
 ID ABU23300 standard; protein; 489 AA.
 XX
 AC ABU23300;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by prokaryotic essential gene #8827.
 XX
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Bordetella pertussis.
 XX
 PN W020027183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zykkind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PDSB; ACA27170.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 51224; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 489 AA;

Query Match 28.8%; Score 690; DB 6; Length 489;
 Best Local Similarity 33.4%; Pred. No. 1e-48;

Matches 161; Conservative 93; Mismatches 200; Indels 26; Gaps 8;

QY 8 YNNGEKKSVNOEITSPIDSSLG---FVPMATREEDVHAKAGREALPMAALTVE 63
 DB 19 YIDGKWKVAGNPSI--PVDNPSSTGKTIISVPLGKREAEQADAEALPMSARTGKE 76
 QY 64 RAQYVHKADIIERDKEETATLAKESKAYNASVTEVRTDLRYAEKGRISTSD 123
 DB 77 RAAILKMKQIMMENQODLAVMTSEOGKVPBAGEIYVYASFEWFAEAKRI---- 131
 QY 124 EGGKMDASTGHTLAIVRRQVGIVLAIAPYNYVNLGSKIAIPALIGVNVWFKPTQGS 183
 DB 122 DDDIIQSPKGGGRMLWVKQPIGVTAITPWNFPAMITRKVPALAACTWYVKRQGP 191
 QY 184 VSGIVLAKAPAPAGLPAGVFNITTGSGISGDIYVEHEVNFINTGSTPVGQRIKLAG 243
 DB 192 LITALLAVLAVLAGVAGVGVVVTGSSREIGALCESEVVRKLSFGSTEVGRTIMEQCA 251
 QY 244 --MRPIMLELGGKMDAGIVLADADLDMNAKOIYAGADVSGGQCTAKRVLVYVEVADELA 301
 DB 252 PTIKCLISTELGNAIPETVPDDADLDAIDGILASKRYRNAQCVCANRIYVQAGYEEVA 311
 QY 302 EKISENVAKLSVGDEPD--NATVTPVIDNSADFIESTLVVDARQKAKEL---NEFKRDR 357
 DB 312 KRLVEKVNAMKVGDEGEQVTOGPLIDTAIVAKVOGHIDATRAHAKVANGEPHALGGS 371
 QY 368 LITGLFDHVTLDMKLAMEEPGPILPIIRVXDABEVAVALANKSDFGLOSFTNDFOKA 417
 DB 372 FPGQFVVRDVTQSMFPAVEETFGPVAPLRFKPSSEDEIVGMANNITFGLAAVFTRDYARI 431
 QY 418 FPIANKLEVTGTHIN-----NKTGRGPDNFPPLGLKSGSAGVQGRIRYSIEANTVKSITLV 472
 DB 432 KWVSEALEYIGVINTGIISNEVG-----PFGVGVSGLGRBGSKYGIEDYIELKCYLV 485
 QY 473 DM 474
 DB 486 DL 487

RESULT 73
 ADS22151
 ID ADS22151 standard; protein; 480 AA.

XX AC
 XX ADS22151;

DT 02-DEC-2004 (first entry)
 XX Bacterial polypeptide #11184.

KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY) CAO Y.

XX (HINK) HINKER G J.

XX (SLAT) SLATER S C.

XX (CHEN) CHEN X.

XX (GOLD) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide

XX for expression of a polynucleotide encoding a polypeptide from a

XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 11184; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for providing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 480 AA;

Query Match 28.7%; Score 687; DB 8; Length 480;
 Best Local Similarity 32.1%; Pred. No. 1.8e-48;

Matches 155; Conservative 105; Mismatches 207; Indels 16; Gaps 6;

QY 1 LTKEQNVNCEWK--SSVNOEITSPIDSSLG--FVPMATREEDVHAKAGREALPMAA 58
 DB 3 MVQEKLPFGFBKDSSTGETFEDINPATYLENLATIQVAGABDVMAVEAABAGFRLWNE 62

QY 59 LTVYERAOYLHKAADIIRERDKEIATVLAKESIKAYNAVSEVYRTADLIRYAAEGIRL 118
DB 63 VPAKAAAEVLFRARLLOERKEELSLVMTKEGKVLPETRGDVQBAIDITNTAAAGEGRM 122
QY 119 ST3ADEGKMDASTGHKLAVIRROPVGIYLAIAFYNYPVNLSSGKIAPALIGNVVMPKP 178
DB 123 L-----GETTSELKKEKFCMTTIRPIGVGLITPMFPAIAPMKWMPALVAGNAVFP 177
QY 179 PTQGVSGVLAKAFAPAGVNTTGRSGEIDYVEHEVNFINTGSTPVQRI 238
DB 178 ASDTPLAFLIEVLSEAGLPPGVINLVTPGCGTVKAVQHPRIKASIFTSGLDTGKWI 237
QY 239 GKLAG--MRPIMEIGGKXAGIYLAADIDDNAKOIVAGAYDSSGORTAIRVVEEV 296
DB 238 MEBCSKTRKRVSLGSKNPVIMVMDADLELAEGLWGAFTTGORTATSLIHEKI 297
QY 297 ADELAEKISENVAKLSVDP--FDNATVTVTVIDNSADFIESTLVVDROKAKELNEFKR- 354
DB 298 KQBFIRKLAKKAKSLIGGGLPEPTDIGVINKAQEKIERVYKIGKEGATILYGNRI 357
QY 355 ----DGRLLTPGLFDHTLDMKLAMEEPPIPIIRYKDAEBAVANKSPDGLQSVF 410
DB 358 DGPLPEYFEPTIFLTVRPMRIAGBEIRGPVIGITVSDLEBALITLANSYKGLSSAIY 417
QY 411 TRDPQKAPFANKLEVTGTHINNKTRGPDNPPFLGKSGAGV--QGIVSTEAMTNVKS 469
DB 418 TGNIGNAFRAIEKVENGITVYVNAPTIGAEVHLPEFGVKGITGNGFREAGTEAVKESEVA 477
QY 470 IVL 472
DB 478 VVI 480

RESULT 74
ADN25659
ID ADN25659 standard; protein; 467 AA.
AC ADN25659;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #8312.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX WPI; 2004-061375/D6.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 8312; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC transforming the recombinant DNA construct and a method of producing a
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content. Improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 467 AA;
Query Match 28.7%; Score 686; DB 8; Length 467;
Best Local Similarity 36.1%; Pred. No. 2,1e-48;
Matches 168; Conservative 91; Mismatches 203; Indels 4; Gaps 4;
QY 8 YNCGEMKSSVNOIELSPIDDSLGFPVPMTEEVDMKAKRELLPAAVLTVYERAY 67
DB 5 WLAGSATGDNELTVNPTGYVATVAVPSSEVDQAAAHGALDAAPALPAHVRATA 64
QY 68 LHKAAADIIRERDKEIATVLAKESIKAYNAVSEVYRTADLIRYAAEGIRLSTADEGK 127
DB 65 LAHVARRLOERSEBELAQILTAESEKPILMARAEVGERAVSFTWMAEB--VRRSGVQRID 123
QY 128 MDASTGHKLAVIRROPVGIYLAIAFYNYPVNLSSGKIAPALIGNVVMPKPPTQGSVGL 187
DB 124 TTSAGEGRIALIRVPKGAIVLAISPFPPLNVAHVAPAIAGVAPIVKPPATPPLSL 183
QY 188 VLAKEFAEAGLPAQVNTTGRSGEIDYVEHEVNFINTGSTPVQGRICKAGMRPI 247
DB 184 VIGELIARLTDLPAGAFSVLP--MPNELAOPLITDERLPVLSFTGG--QFGWELKMKAPNKHV 241
QY 248 MLELGKDGAGIYLAADIDDNAKOIVAGAYDSSGORTAIRVVEEVADELAEKISEN 307
DB 242 TLELGNAAAVTLADIDDMAARRIALRGNNQAGVCIGQVITYEDAYDEFAKLTVR 301
QY 308 VAKLSVGPFDNAT--VTPVIDNSADFIESTLVVDAROKAKELNEFKRDLITPGLFDH 366
DB 302 VELVLTGPADBATOVGPLINEAARAVEQWNEVAGGATVLTGTGRDGVMTAPTVLAD 361
QY 367 VTLDMKLAMEEPPIPIIRYKDAEBAVANKSPDGLQSSVFPTRDPKADIANKLEV 426
DB 362 VPDDAKVCEEVFGVMMVLRVSGVDEFAEVNNSKYGIGAGVFTRLDALTAFRAHRELKV 421
QY 427 GTVHINNKTRGPDNPPFLGKSGAGVQGIIRYSIEAMTNVSIYL 472
DB 422 GGVIIIGDVPYTRADQMPYGVKSGVKGEGVRSAMDLLDERVMVL 467

RESULT 75
ADN24384
ID ADN24384 standard; protein; 458 AA.
AC ADN24384;

XX 02-DEC-2004 (first entry)
DT Bacterial polypeptide #7037.
DE
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.
XX Bacteria.
OS
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 7037; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 458 AA;

Query Match 28.5%; Score 682.5; DB 8; Length 458;
Best Local Similarity 36.1%; Pred. No. 3.9e-48;
Matches 168; Conservative 84; Mismatches 199; Indels 15; Gaps 7;

QY 11 GEMKSSVNOIEILSPIDSSSLGFVPMATREBYVDHAKAGREALPMALTTVERAQYLHK 70
DB 1 GERVTRPRTLDVDFYTGTVGTPLAS---VDVRAAFAEYAAAYOALTRYERSQILER 57

QY 71 AADIIEERDKERIAIVLAKELSKAVNASYTEVVRATDLIRYAABEGIRLSTADEGKMDA 130
DB 58 AASIRBERTEEASDLISLESGLSKQDSRYEIGRVADVLFKFASTIALR-----DQGQSFSC 112
QY 131 ST---GKTLAVI-RRQV-GIVLAIAPYNYVNLSSGKIAPALIGNVVMFKPPTQSSVS 185
DB 113 DLTPGKKRRYFSSQREPLAGVIVATPPNHMNOVAHKIAPAIATNNRVLLKPSKXPLS 172
QY 186 GLVLAKFAEAGLPAQVNTITGSGSEIGDYIVHEEBVNFNFPGSTPVGGRIGKLGM 245
DB 173 ALYLADVLYEAGLPAPMLQVLTGDPRETADLITPLAELVTFQGVATIGKTIAPARA 232
QY 246 PIMLELGGKADGIVLADADLDNAAKQIVAGAYDYSQGRCTAIRKVLVVEEYADELAEKIS 305
DB 233 RVLELGNDPLIYLDADLERAAATLAVQSGYSKNSGKRCITAVKMLVQKSVAADPTDLVV 292
QY 306 ENVAKLSVGDPFD-NATVTPYIDNSADFTISLVVDARQKAKELNEKRGRLITPOLF 364
DB 293 EKTRAWTFGDPFDASSQWGTIVDVAALQLEFARVMEBAVASGARLLTGQRNGALYAPTVL 352
QY 365 DHVTLDMKLAWEPPGPILPTIIRVKDAEBAVAIAINKSDPGLQSSVETRDPOKAPDIANKL 424
DB 353 DGVDPSMTLYREBETFGVPSPILITFTDLDARISNGTAFGLSSGVCTNRQDAITRFINEL 412
QY 425 EVGTVHINKTGRGPDNPFLGLKSGAGV-QGIRYSIEAMTNVYS 469
DB 413 RVGTINWEVPGYRIELTFPGIGKDSGLGYKEGVQAMKSFNLT 458

Search completed: August 28, 2005, 10:24:46
Job time : 177 secs

1 LENGTH: 474
 2 TYPE: PRT
 3 ORGANISM: Streptococcus pneumoniae
 4 FEATURE:
 5 OTHER INFORMATION: glyceraldehyde-3-phosphate dehydrogenase, NADP-dependent (gapN)
 6 OTHER INFORMATION: Cellular location: cytoplasm
 7 OTHER INFORMATION: Similar to strain R6 sequence 15903072 (0.E+01)
 8 US-10-472-928-2228
 9
 10 Query Match 75.7%, Score 1812, DB 17, Length 474,
 11 Best Local Similarity 74.1%, Pred. No. 9, 1e-137,
 12 Matches 349, Conservative 57, Mismatches 65, Indels 0, Gaps 0,
 13
 14 QY YQNVNCGEKKSSVNOQIEILSPIDDSLGVPVPMTRREVDHAKAGREALPAAALTVYER 64
 15 DQ YQNVNCGEKKSSVNOQIEILSPIDDSLGVPVPMTRREVDHAKAGREALPAAALTVYER 63
 16 4 YQNVNCGEKKSSVNOQIEILSPIDDSLGVPVPMTRREVDHAKAGREALPAAALTVYER 63
 17 QY AQLVHKRAADIIEEDKEELATVLAKEISKAYNMSVTEVPTADLIRYAAEEGRLSTADE 124
 18 DQ AQLVHKRAADIIEEDKEELATVLAKEISKAYNMSVTEVPTADLIRYAAEEGRLSTADE 123
 19 64 AQLVHKRAADIIEEDKEELATVLAKEISKAYNMSVTEVPTADLIRYAAEEGRLSTADE 123
 20 QY GGRKDDSTGHKLAVIRRPQVGIYLAIAPNYPVNLSGSKIAPALIGNVNMFKEPTQGSV 184
 21 DQ GGRKDDSTGHKLAVIRRPQVGIYLAIAPNYPVNLSGSKIAPALIGNVNMFKEPTQGSV 183
 22 124 GGRKDDSTGHKLAVIRRPQVGIYLAIAPNYPVNLSGSKIAPALIGNVNMFKEPTQGSV 183
 23 QY SGVLAKAPAEALPAGVFRNTITGRGSEIGDYIVHEEYVNFNFGSTVVGQRIKGLAM 244
 24 DQ SGVLAKAPAEALPAGVFRNTITGRGSEIGDYIVHEEYVNFNFGSTVVGQRIKGLAM 243
 25 184 SGVLAKAPAEALPAGVFRNTITGRGSEIGDYIVHEEYVNFNFGSTVVGQRIKGLAM 243
 26 QY RPLMLELGGDAGIVLADADLDNAAKOIVAGADVSGQRCTAIKRVLYVEEVADELAEXI 304
 27 DQ RPLMLELGGDAGIVLADADLDNAAKOIVAGADVSGQRCTAIKRVLYVEEVADELAEXI 303
 28 244 RPLMLELGGDAGIVLADADLDNAAKOIVAGADVSGQRCTAIKRVLYVEEVADELAEXI 303
 29 QY SENVAKLSVGDPPDNATVTPVVIDNSADPIESLVDARQKAKELNEFRDRGLLTPGLF 364
 30 DQ SENVAKLSVGDPPDNATVTPVVIDNSADPIESLVDARQKAKELNEFRDRGLLTPGLF 363
 31 304 SENVAKLSVGDPPDNATVTPVVIDNSADPIESLVDARQKAKELNEFRDRGLLTPGLF 363
 32 QY DHTYLDMLKAMEEPPGPIPLIRVYKDAEBAVALANKSDGLOSSTVTRPQQAFLIANKL 424
 33 DQ DHTYLDMLKAMEEPPGPIPLIRVYKDAEBAVALANKSDGLOSSTVTRPQQAFLIANKL 423
 34 364 DHTYLDMLKAMEEPPGPIPLIRVYKDAEBAVALANKSDGLOSSTVTRPQQAFLIANKL 423
 35 QY EVGVVHHNNNGRGPDNPPLGLKGSAGVQIGRYSIEAMTVKSLVLDPMK 475
 36 DQ EVGVVHHNNNGRGPDNPPLGLKGSAGVQIGRYSIEAMTVKSLVLDPMK 474
 37 424 EVGVVHHNNNGRGPDNPPLGLKGSAGVQIGRYSIEAMTVKSLVLDPMK 474
 38
 39 RESULT 3
 40 US-10-282-122A-73974
 41 Sequence 73974, Application US/10282122A
 42 Publication No. US20040029129A1
 43 GENERAL INFORMATION:
 44 APPLICANT: Wang, Liangsu
 45 APPLICANT: Zamudio, Carlos
 46 APPLICANT: Malone, Cheryl
 47 APPLICANT: Haselbeck, Robert
 48 APPLICANT: Ohlsen, Kari
 49 APPLICANT: Zyskind, Judith
 50 APPLICANT: Wall, Daniel
 51 APPLICANT: Twilick, John
 52 APPLICANT: Carr, Grant
 53 APPLICANT: Yamamoto, Robert
 54 APPLICANT: Forsyth, R.
 55 APPLICANT: Xu, H.
 56 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 57 FILE REFERENCE: EPIPA 034A
 58 CURRENT APPLICATION NUMBER: US/10/282,122A
 59 CURRENT FILING DATE: 2003-02-20
 60 PRIOR APPLICATION NUMBER: 60/191,078
 61 PRIOR FILING DATE: 2000-03-21
 62 PRIOR APPLICATION NUMBER: 60/206,848
 63 PRIOR FILING DATE: 2000-05-23
 64 PRIOR APPLICATION NUMBER: 60/207,727
 65 PRIOR FILING DATE: 2000-05-26

```
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 73974
/ LENGTH: 474
/ TYPE: PRF
/ ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73974

Query Match          75.3%; Score 1802; DB 15; Length 474;
Best Local Similarity 73.7%; Pred. No. 5 8e-136;
Matches 347; Conservative 58; Mismatches 66; Indels 0; Gaps 0;

QY 5 YONVYNGEMKSSVNOIELSPIDSSLGFPVPMTRREVDHMKAGREALPAWALTYVER 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 4 YONLVNGKMKSSBOETITVSPINOELGTVPAMTQTEADEAMQARALPAWALSAIR 63
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 65 AOVHKAADIIRDEKEIATVLAKEISKAYNASVTVTADLIRAAEGIRLSADE 124
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 64 AAVHKAADIIRDEKEIATVLAKEISKAYNAKIGVVTADLIRAAEGIRLSADE 123
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 125 GSKMDASTGHKLAVIRROPVGIYLAIAFYNPVNLGSKIAAPALIGNVMFAPPTQGSV 184
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 124 GGGFPAASKNKLAVVREVGIVLAIAPFNPVNLGSKIAAPALIGNVMLKPTQGS 183
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 185 SGVLAKAFPAEAGLPAGVFNITIGRSGEIGDYIYEHEEVNFTNFTSGTPEVGORIGLQAM 244
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 184 SGVLAKAFPAEAGLPAGVFNITIGRSGEIGDYIYEHEEVNFTNFTSGTPEVGORIGLQAM 243
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 245 RPIHMEELGKDGAGIVLADADLNNAAQIYAGAYDYSGORCTAIKRYLVVEEVADELAEKI 304
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 244 RPIHMEELGKDGAGIVLADADLNNAAQIYAGAYDYSGORCTAIKRYLVVEEVADELAEKI 303
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 305 SENVAKLSVGDPPDNATVTVPIDDNGADFTESLVVDAROKGAKELNEFRDGRILTRPGF 364
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 304 QEBVSKLIVGDPFNDADITVVIDNASADFTWGLIEDAOKERQALTPIRREGNLPVLF 363
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 365 DHVTLDMKLAMEBPPGPIPIIRVKDAEBAVAIALNKSDFGLOSSVFTTRDFOKAFDIANL 424
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 364 DQVTKDMKXAMEBPPGPIPIIRVASVEBALIAPANSSEFGLOSSVFTTRDFOKAFDIANL 423
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 425 EYGTVINNKTRGPGNPFPLGKSGAGVQGIIRYSIEAMTVKSIIVLDMK 475
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 424 EYGTVINNKTRGPGNPFPLGKSGAGVQGIIRYSIEAMTVKSIIVLDMK 474
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
US-10-282-122A-72316
/ Sequence 72316; Application US/10282122A
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
```

```
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT FILING DATE: US/10/282,122A
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 72316
/ LENGTH: 475
/ TYPE: PRF
/ ORGANISM: Streptococcus mutans
US-10-282-122A-72316

Query Match          71.0%; Score 1700; DB 15; Length 475;
Best Local Similarity 66.9%; Pred. No. 9.1e-128;
Matches 318; Conservative 75; Mismatches 82; Indels 0; Gaps 0;

QY 1 LKREYONVNGEMKSSVNOIELSPIDSSLGFPVPMTRREVDHMKAGREALPAWALTYVER 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MKRQYINVNGEMKSSVNOIELSPIDSSLGFPVPMTRREVDHMKAGREALPAWALTYVER 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 YVERAQLHKAADIIRDEKEIATVLAKEISKAYNASVTVTADLIRAAEGIRLSADE 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 YVERAQLHKAADIIRDEKEIATVLAKEISKAYNASVTVTADLIRAAEGIRLSADE 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 SHDEGKMDASTGHKLAVIRROPVGIYLAIAFYNPVNLGSKIAAPALIGNVMLKPTQGS 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 121 SHDEGKMDASTGHKLAVIRROPVGIYLAIAFYNPVNLGSKIAAPALIGNVMLKPTQGS 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 OGSVSGIVLAKAFPAEAGLPAGVFNITIGRSGEIGDYIYEHEEVNFTNFTSGTPEVGORIGL 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 181 OGSVSGIVLAKAFPAEAGLPAGVFNITIGRSGEIGDYIYEHEEVNFTNFTSGTPEVGORIGL 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 LAGMPRIMLELGKDGAGIVLADADLNNAAQIYAGAYDYSGORCTAIKRYLVVEEVADEL 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 241 LAGMPRIMLELGKDGAGIVLADADLNNAAQIYAGAYDYSGORCTAIKRYLVVEEVADEL 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 AEKISENVAKLSVGDPPDNATVTVPIDDNGADFTESLVVDAROKGAKELNEFRDGRILTR 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 301 AEKISENVAKLSVGDPPDNATVTVPIDDNGADFTESLVVDAROKGAKELNEFRDGRILTR 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 PGLPHVTLDMKLAMEBPPGPIPIIRVKDAEBAVAIALNKSDFGLOSSVFTTRDFOKAFDI 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 361 PGLPHVTLDMKLAMEBPPGPIPIIRVKDAEBAVAIALNKSDFGLOSSVFTTRDFOKAFDI 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 421 ANKLEVGTVHINKTRGPGNPFPLGKSGAGVQGIIRYSIEAMTVKSIIVLDMK 475
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 421 ANKLEVGTVHINKTRGPGNPFPLGKSGAGVQGIIRYSIEAMTVKSIIVLDMK 475
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 5
US-09-815-242-13316
Sequence 13316, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haseibeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13316
LENGTH: 439
TYPE: PRF
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13316

Query Match 71.0%; Score 1699; DB 9; Length 439;
Best Local Similarity 74.5%; Pred. No. 9,7e-128;
Matches 327; Conservative 55; Mismatches 57; Indels 0; Gaps 0;

QY 37 MTRREVDHAKGREALPMAALTYVERAOYLHKADIERDKBEIATVLAKESKAYNA 96
DB 1 MTOLEDEMOARALPMPWRALSAIERAAVYHKTAIERDKKEIGTTLAKESKAYNA 60
QY 97 SVTETVTRADLIRYAAEEGIRLSTSDSEGKMDASTGHLAVIRROPVGIVALAPYNY 156
DB 61 AIGGVTRADLIRYAAEEGLRTGQAMEGGFEAASKNKLAVRRPVGIVLAIAPNVP 120
QY 157 VNLSSKIPALIGNVVFKPPTQGSVGLYLAKFAEAGLPAGVFNTITGRGSEIGDY 216
DB 121 VNLSSKIPALIGNVVFKPPTQGSISGLILAKAFEEAGIPAGVFNTITGRGSEIGDY 180
QY 217 IYEHHEVNFINTGSPVGORIGKLAGMPPIMELEGKAGIVLADADLNNAKOIVAA 276
DB 181 IYEHHEVNFINTGSPVGORIGKLAGMPPIMELEGKAGIVLADADLNNAKOIVAA 240
QY 277 YDYGSGORTAIRVAVVEEVADELAEKISENVAKLSVGPFDNATVPYIDNSADFI 336
DB 241 FYSGSGORTAIRVAVVEEVADELAEKISENVAKLSVGPFDNATVPYIDNSADFI 300
QY 337 LVVDAROKAKELNEFKRQDGLTLTGLPHVTLDMKLANEBEPGPILPIIRVAKDEAVA 396
DB 301 LLEDQOEKKAQALTPIKREGNLMVPLFDQVTKDMKVAEBEPGPVLPPIIRVASEAIA 360
QY 397 IANKSPDFGSSVFTRDPOKAFDIAKLVGTWTHINKKRGPDNFPPIGLGSGAGVYG 456
DB 361 PANESBFGLOSSVFTNDPKAFBIAEKLEVGIVHINKKRGPDNFPPIGLGSGAGVYG 420

QY 457 IRYSTIEMNTVKSIVLDMK 475
DB 421 IKYSIEMNTVKSIVFDVK 439

RESULT 6
US-10-617-320-3804
Sequence 3804, Application US/10617320
Publication No. US20050136404A1
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/617,320
FILING DATE: 10-Jul-2003
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3804:
SEQUENCE CHARACTERISTICS:
LENGTH: 436 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...436
SEQUENCE DESCRIPTION: SEQ ID NO: 3804:
US-10-617-320-3804

Query Match 70.7%; Score 1692; DB 18; Length 436;
Best Local Similarity 74.9%; Pred. No. 3.5e-127;
Matches 326; Conservative 53; Mismatches 56; Indels 0; Gaps 0;

QY 41 EYDHAHMKGREALPMAALTYVERAOYLHKADIERDKBEIATVLAKESKAYNA 100
DB 2 EADBSGQARALPMPWRALSAIERAAVYHKTAIERDKKEIGTTLAKESKAYNA 61
QY 101 VTRTADLIRYAAEEGIRLSTSDSEGKMDASTGHLAVIRROPVGIVALAPYNYVNL 160
DB 62 VTRTADLIRYAAEEGLRTGQAMEGGFEAASKNKLAVRRPVGIVLAIAPNVPVNL 121
QY 161 GSKIPALIGNVVFKPPTQGSVGLYLAKFAEAGLPAGVFNTITGRGSEIGDYIYEH 220

Db 122 ASKIAFALLAGNVVMEKPTQSSISGLLAKAFEEAGIPAGVFNITITGSEIGDIYIH 181
Qy 221 EEVNFNFTGSIPTVGORIGKLAQMRPIMELGSKDAGIVLADLDNNAKQIVAGAYDYS 280
Db 182 KEVNFNFTGSIPTIGRIGRLAGMRPIMELGSKDALVLEDADLEHAAKQIVAGAFSYS 241
Qy 281 GORCTAIKRVLYVEEVADELAEKISNVAKLSVGDPPDNATPTPVVDNSADIEELVVD 340
Db 242 GORCTAIKRVLYVESVADKATLLQSEVSCLTVGDPPDNADITPVVDNSADIEELVVD 301
Qy 341 AROKGAKEINEFEKDRRLTPGLFDHVTLLDMKLAEEEPGPIIPIRVKDAEEAVALANK 400
Db 302 AORKEAQUALTPIRKREBNLMPVFPDQYTKOMKVAEPEPGPVPIIRVAVSEALIAFANE 361
Qy 401 SDGLOSSTFTDPOKAFDIANKLEVGVTHINNKTRGDPNPPFLGKSGAGVQIGIRYS 460
Db 362 SERGLSSVFTNDFKAFETAEKLEVGVTHINNKTRGDPNPPFLGKSGAGVQIGIRYS 421
Qy 461 IEAMTVKSIIVLDMK 475
Db 422 IEAMTVKSIIVFDVK 436

RESULT 7
US-10-282-122A-74625
Sequence 74625, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remainder of prior application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 74625
LENGTH: 475
TYPE: PRT
ORGANISM: Streptococcus pyogenes
US-10-282-122A-74625

Query Match 68.5%; Score 1639; DB 15; Length 475;
Best Local Similarity 65.5%; Pred. No. 7.2e-123;
Matches 311; Conservative 77; Mismatches 87; Indels 0; Gaps 0;
Qy 1 LTRKGVYONVNGEMKSVNOIEILSPIDBSLGFVPMATREEDVHAMKAGREALPAAAL 60
Db 1 MAKQYGNLVNGEMKSVSENIITTPATGTEELSSVPMPTQAEVDAYASAKKLSDBRAL 60
Qy 61 VYERAPYLKRAADIIERDKESIATYIAKEISAYNNAVTEVVRTPADLLIYAAEEGRLST 120
Db 61 VYERAPYLKRAADIIERDKESIATYIAKEISAYNNAVTEVVRTPADLLIYAAEEGRLST 120
Qy 121 SADEGSKMASYGHKLAIVRRQPVGIVLAIAVYNPVNLGSKIAFALLAGNVVMEKPT 180
Db 121 EYLEGSGFPAASKKRIATVRRPVGIVLAISPNNPVNLGSKIAFALLAGNVVMEKPT 180
Qy 181 QGSVSGLVLAFAKFAEAGLIPAGVFNITITGSGSEIGDIYVEHEEYFNFTGSIPTVGORIGK 240
Db 181 QGSISGLLAEAFEAEGIPAGVFNITITGSGVIGDIYVEHEEYFNFTGSIPTVGORIGK 240
Qy 241 LAGMRPIMELGSKDAGIVLADLDNNAKQIVAGAYDYSGORCTAIKRVLYVEEVADEL 300
Db 241 LAGMRPIMELGSKDAGIVLADLDNNAKQIVAGAYDYSGORCTAIKRVLYVEEVADEL 300
Qy 301 AEKISENVAKLSVGDPPDNATPTPVVDNSADIEELVVDARQKAEINEFEKDRRLT 360
Db 301 AEIKTLVKLSVGMPEDDADITPLDTSAAOFBEILDADVDKGTALTAFRREBNLIS 360
Qy 361 PGLFDHVTLLDMKLAEEEPGPIIPIRVKDAEEAVALANKSDGLOSSTFTDPOKAFDI 420
Db 361 PVLFDHVTLLDMKLAEEEPGPIIPIRVVTEBAIKISNESBYGLQASLFTTNFPARFI 420
Qy 421 ANKLEVGVTHINNKTRGDPNPPFLGKSGAGVQIGIRYSIAMTVKSIIVLDMK 475
Db 421 AEGLEVGVTHINNKTRGDPNPPFLGAKKSGAGVQVKSIAMTVKSIIVFDIQ 475

RESULT 8
US-10-282-122A-53543
Sequence 53543, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931


```
QY 63 ERAQYLHRAADIIEBDEKEIATVLAKEISKAYNASVTEVTRADLIRYAABEGIRLSTSA 122
DB 78 KBAELHKAALILKEHKAPIAECLVEIKAKPAKDAVEVRSQDLVSYAABEGVRLIGE- 136
QY 123 DEGGKM---DASTG---HKLAVIRROPVGIIVLAIAPYNPVNLGSGKIPALIGGNVWF 176
DB 137 ---GKLLVSDSPFGNERNKCYCLSSKIPGLGVLAIPFPNTPVNLAVSKIGIPALINAGNALV 193
QY 177 KPPTQSVSGVLVLAAPABAGLPAGVFNITITGSGSEIGDYIYEHEEVNFINTGSTPVQ 236
DB 194 KPPTQGAVALHMHVHCFHLAGFPKGLISCVTGKSEIGDFLTMHPGVNCSIFGTG- DTGI 252
QY 237 RIGKLAKMRPIMLEIGCKDAGIVLADADLDNNAKOIVAGAYDYSGORCTAIRVLVEEV 296
DB 253 AISKAGAVPLOMEIGCKDAGIVLADADLDVLANIVKGFYSQGRCTAIRVAVVIMESI 312
QY 297 ADELAEKISENVAKLSVGDPPFNATVTVYIDNSADFIESLVVDAROKGAKELNEPKRD 356
DB 313 ADAVVOQVNAKLAKLVGPPEDSDITTPVTESSANFIEGLVMDAKKGAFCOEYRREG 372
QY 357 RLITPGLFDHVTLDKMLAMEBPPGPIPLPIIRVKDAEAVANANKSDPGLQSSVFTFPOK 416
DB 373 NLIMPLLDHYVAPDMRIAMEBPPGPIPLPIIRINSVEEIGHHCNANSPGLQGCIFTRDINK 432
QY 417 AFDIANKLEVGTVHINNKTGRGPDNFPPLGLKSGAGVQIGRYSIEAMTNVKSIVLDM 474
DB 433 ALLSDAMETGTVOINSAPARGPDHFPQGLKQSGIGSOGITNSIMMTKXKSTVINTL 490
```

RESULT 15

```
US-10-437-963-189014
; Sequence 189014, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 189014
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85563C.1.pep
US-10-437-963-189014
```

```
Query Match 47.9%; Score 1146.5; DB 16; Length 503;
Best Local Similarity 48.5%; Pred. No. 2.9e-83;
Matches 234; Conservative 87; Mismatches 144; Indels 17; Gaps 6;

QY 5 YONYNGEMKSVN--QIEILSPIDSSIGFVPAMTRREVDHAMKAGREALPAMALTVY 62
DB 19 YKYVADGEMKSVASGSVAIINPTTRKTQYKVOACSGEENKVMETAKVAKAMARTPLM 78
QY 63 ERAQYLHRAADIIEBDEKEIATVLAKEISKAYNASVTEVTRADLIRYAABEGIRLSTSA 122
DB 79 KBAELHKAALILKEHKAPIAECLVEIKAKPAKDAISEVRSQDLVSYAABEGVRLIGE- 137
QY 123 DEGGKM---DASTG---HKLAVIRROPVGIIVLAIAPYNPVNLGSGKIPALIGGNVWF 176
DB 138 ---GKLLVSDSPFGNERNKCYCLSSKIPGLGVLAIPFPNTPVNLAVSKIGIPALINAGNALV 194
```

```
QY 177 KPPTQSVSGVLVLAAPABAGLPAGVFNITITGSGSEIGDYIYEHEEVNFINTGSTPVQ 236
DB 195 KPPTQGAVALHMHVHCFHLAGFPKGLINCVTGKSEIGDFLTMHPGVNCSIFGTG- DTGI 253
QY 237 RIGKLAKMRPIMLEIGCKDAGIVLADADLDNNAKOIVAGAYDYSGORCTAIRVLVEEV 296
DB 254 AISKAGAVPLOMEIGCKDAGIVLADADLDVLANIVKGFYSQGRCTAIRVAVVIMESI 313
QY 297 ADELAEKISENVAKLSVGDPPFNATVTVYIDNSADFIESLVVDAROKGAKELNEPKRD 356
DB 314 ADIVEKIVAKLAKLVGPPEDSDITTPVTESSANFIEGLVMDAKKGAFCOEYRREG 373
QY 357 RLITPGLFDHVTLDKMLAMEBPPGPIPLPIIRVKDAEAVANANKSDPGLQSSVFTFPOK 412
DB 374 NLIMPLLDHYVAPDMRIAMEBPPGPIPLPIIRINSVEEIGHHCNANSPGLQGCIFTRDINK 433
QY 413 DFOKAFDIANKLEVGTVHINNKTGRGPDNFPPLGLKSGAGVQIGRYSIEAMTNVKSIVL 472
DB 434 DINKAMISDAMETGTVOINSAPARGPDHFPQGLKQSGIGSOGITNSIMMTKXKSTVINTL 493
QY 473 DM 474
DB 494 NL 495
```

RESULT 16

```
US-10-424-599-187423
; Sequence 187423, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 187423
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(503)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_140255C.1.pep
US-10-424-599-187423
```

```
Query Match 43.4%; Score 1038.5; DB 15; Length 503;
Best Local Similarity 46.2%; Pred. No. 1.4e-74;
Matches 223; Conservative 80; Mismatches 163; Indels 17; Gaps 7;

QY 5 YONYNGEMKSVN--QIEILSPIDSSIGFVPAMTRREVDHAMKAGREALPAMALTVY 62
DB 17 YKYVADGEMKSVASGSVAIINPTTRKTQYKVOACSGEENKVMKMDLAKSAQGLAKMTPLM 76
QY 63 ERAQYLHRAADIIEBDEKEIATVLAKEISKAYNASVTEVTRADLIRYAABEGIRLSTSA 122
DB 77 KBAELHKAALILKEHKAPIAECLVEIKAKPAKAVTEVRSQDLVSYAABEGVRLIGE- 133
QY 123 DEGGKM---DASTG---HKLAVIRROPVGIIVLAIAPYNPVNLGSGKIPALIGGNVWF 177
DB 134 GEGKVLVSDSPFGNERNKCYCLSKIPGLGVLAIPFPNTPVNLAVSKIPALINAGNALV 193
QY 178 KPPTQSVSGVLVLAAPABAGLPAGVFNITITGSGSEIGDYIYEHEEVNFINTGSTPVQ 237
DB 194 PPTQGAVALHMHVHCFHLAGFPKGLINCVTGKSEIGDFLTMHPGVNCSIFGTG- DTGI 252
QY 238 IGLAKMRPIMLEIGCKDAGIVLADADLDNNAKOIVAGAYDYSGORCTAIRVLVEEVA 297
```

Db 253 ISKKGMIPLQWELGSKACIVLEISDLQCMKMTSYLNKQPNPGRCTAVKVLVWESKA 312
QY 298 DELAKISENVAKLSVGDPEFNATVTPVINDNSADFTESLVYDAQKQAK-----ELNMF 352
Db 313 DALVEKAKVAKVAKLVGPEPDDCDITPVSESSANFIEGLVDAKEKATFPMGACEXK 372
QY 353 KRDGRLTPGLFDHYTLDMKLAWEPPFPIIPRIYKQAE-AVALANKSDFGLQSVFT 411
Db 373 KREGNPIWFLLDXVRPDIIXIMBEPFVLPVIRINSVEEXGIHHCNANSFGLDMSPX 432
QY 412 RDPQAFDIANKLEVTGTHINNKTRGPNPPLGLKSGAGVQGISIEAMTVKSTIV 471
Db 433 RQVNFAMISDMETGTVOINSAPARGPDPHFGQIKDSIGSGIGITSINMTKVKTV 492
QY 472 LDM 474
Db 493 INL 495

RESULT 17

US-10-424-599-187419
; Sequence 187419, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 187419
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1):(434)
; FEATURE:
; OTHER INFORMATION: unsure at all Xaa locations
US-10-424-599-187419

Query Match 41.9%; Score 1001.5; DB 15; Length 434;
Best Local Similarity 48.6%; Pred. No. 1e-71;
Matches 204; Conservative 73; Mismatches 132; Indels 11; Gaps 5;

QY 5 YQNYNGEKKSSVN--QIEILSPIDDSLGFPVPMREVDHAKMGREALPAMALTYE 62
Db 17 YKTYADGVWKKASGKSVSTINFTTRKTQYVQASQEBKVNKMDLAKSAQKLMATPLM 76
QY 63 EBAQYVHKADIIERDKIEIATVLAKEISKAVNASVTEVVRTADLIRYAEISIRLSTSA 122
Db 77 KRAELHKAALIKKEKPTIACGLVKEIKPAKDAWMEVRSGLDVSTYAEISVRL---L 133
QY 123 DEGGKM--DASTGH--KLAIVIRROQVGIYLAIAPNYVNVNLSGSKIAPALIGSNVMEK 177
Db 134 GEGKFLVSDSPGNERTKCYLTSKIPLVGIYXAIIPPYVFNLAIVKKIAPALIAINSIVLK 193
QY 178 PPTQGSVSGVLAKAFABAGLPAQVFNITGRGSEIGDYIVHEEYVNFNFTGSPVQGR 237
Db 194 PPTQGSVSGVLAKAFABAGLPAQVFNITGRGSEIGDYIVHEEYVNFNFTGSPVQGR 252
QY 238 ICKLAGRPIMLDELGSKDAGIYLAADADLDMNAKQIVAGAYDYSQGRCTAIKEVLVVEEVA 297
Db 253 ISKKGMIPLQWELGSKACIVLEISDLQCMKMTSYLNKQPNPGRCTAVKVLVWESKA 312
QY 298 DELAKISENVAKLSVGDPEFNATVTPVINDNSADFTESLVYDAQKQAKELNEFRDGR 357

Db 313 DALVEKAKVAKVAKLVGPEPDDCDITPVSESSANFIEGLVDAKEKATFPMGACEXK 372
QY 353 KRDGRLTPGLFDHYTLDMKLAWEPPFPIIPRIYKQAE-AVALANKSDFGLQSVFT 411
Db 373 KREGNPIWFLLDXVRPDIIXIMBEPFVLPVIRINSVEEXGIHHCNANSFGLDMSPX 432

RESULT 18

US-10-282-122A-76821
; Sequence 76821, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jiansu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zysek, Daniel
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76821
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Ureaplasma urealyticum
US-10-282-122A-76821

Query Match 36.3%; Score 868.5; DB 15; Length 475;
Best Local Similarity 39.7%; Pred. No. 5.6e-61;
Matches 190; Conservative 95; Mismatches 177; Indels 17; Gaps 5;

QY 5 YQNYNGEKKSSVNQIEILSPIDDSLGFPVPMRE--EVDHAKMGREALPAMALTYE 63
Db 3 YKTYLNGAFVDAKEKELPYVNSNNQIIVVNIHNEININTFEBAHIAVLKFKOTPIKY 62
QY 64 RAQYVHKADIIERDKIEIATVLAKEISKAVNASVTEVVRTADLIRYAEISIRLSTSA- 122
Db 63 RCDLLKLADLDEKQKLAQIISTEIAKGLKDSIEVERADYLRFTVFEYQKLMQKPI 122
QY 123 --DEGGKMDASTGH--KLAIVIRROQVGIYLAIAPNYVNVNLSGSKIAPALIGSNVMEK 176
Db 123 IFDE-----TVHVKAKVATYVYRIIPGVVLAICPFYVPIINLLISKIAPALVSGNSLVY 175


```

Db      61 RLRKADLIEREDELALLETLDGKPLAEAKGDEVEGALDEIRYVAGMARKMGERR 120
Qy      116 I--RLSTADDEGKMDASTGHKLAIVRRQPVIGVLAIAPIYNYVNISSGKIAPALIGNV 173
Db      121 VIPSATIDDE-----ELNYYTRREPLGVGVISPMNPFLLALMWLADALAAAGNT 170
Qy      174 VMFKPPTQGSVSGVLAKAFAGAG---LPAGVFNTITGRSGELIGDYVHEEVPINFTNG 230
Db      171 VVAKSEBQTPPLTALLAELEBAGANNLPKGVVNVVPGFAEVGQALLSPDIDKISFTG 230
Qy      231 STPVGORIGKLA---GMRPIMLELGKDGAGIVLADADLDNAKQIVAGADYSGORCTAI 287
Db      231 STEVGKLIMEAANAANKUKVTLELGKSPYIVFDADLDKAVERYIVFGAFGNAQVCIAP 290
Qy      288 KRVLVVEVADLEAEKISENVAKLS-VGDPFNAT--VTPVIDNSAD-FIESLVVDARQ 343
Db      291 SRLVHESIYDEFEVEKTKERVKKLIGDPLDSDTNVYGPLISEQGFDRVLMGYIEDGKE 350
Qy      344 KGAK-----ELNFKDGRLLTPGLPHTLTLMKLMEEPPGPIPIIRVKDAEAV 395
Db      351 EGAKVLCGERDSKKEYLGGGYVQPTIFDVTPEMKIMKEELFGVLPITIKFKDIDEAI 410
Qy      396 ATANKSDPGLQSSVFTRD-FOKAFDIANKLEVGTVAHNNKTRGSPD-NPPFLGLK-GSGA 452
Db      411 ELANDTEYGLAAVYFTKDLIARAFVAKKALEAGIVVNDVCVAABEPOLPFGGVKSSGI 470
Qy      453 GVO-GIRYSIEAMTVKSVL 472
Db      471 GREHGKYGLEBYTEIKVTI 491

```

RESULT 23

```

; US-10-776-871-21
; Sequence 21, Application US/10776871
; Publication No. US20040132087A1
; GENERAL INFORMATION:
; APPLICANT: Gluckemann, Maria Alexandra
; APPLICANT: Meyers, Rachel
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: NOVEL HUMAN ENZYME FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-193001
; CURRENT APPLICATION NUMBER: US/10/776, 871
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: US/10/175, 696
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 10/067, 668
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/266, 140
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/823, 901
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: PCT/US01/10720
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/193, 920
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/862, 658
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16380
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/205, 675
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/882, 837
; PRIOR FILING DATE: 2001-06-15
; Remaining prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence

```

US-10-776-871-21

Query Match 32.5%; Score 777; DB 16; Length 493;

Best Local Similarity 38.5%; Pred. No. 1,3e-53;

Matches 193; Conservative 92; Mismatches 166; Indels 50; Gaps 17;

```

Qy      12 EWKSSVN--QITLSPIDDS--LGFVPMATREVDHAMAGREAU---PAMALITYERA 65
Db      1 EWVDSASGKTFEVVNPANKEVIGRVPEATAEDVDAVAAEAFKSGFPMKAVPASERA 60
Qy      66 QYLHKAADIERDKESIAVLAKESK--AYNASYTEVVRTADLRYYA-----EBG 115
Db      61 RLRKADLIEREDELALLETLDGKPLAEAKGDEVEGALDEIRYVAGMARKMGERR 120
Qy      116 I--RLSTADDEGKMDASTGHKLAIVRRQPVIGVLAIAPIYNYVNISSGKIAPALIGNV 173
Db      121 VIPSATIDDE-----ELNYYTRREPLGVGVISPMNPFLLALMWLADALAAAGNT 170
Qy      174 VMFKPPTQGSVSGVLAKAFAGAG---LPAGVFNTITGRSGELIGDYVHEEVPINFTNG 230
Db      171 VVAKSEBQTPPLTALLAELEBAGANNLPKGVVNVVPGFAEVGQALLSPDIDKISFTG 230
Qy      231 STPVGORIGKLA---GMRPIMLELGKDGAGIVLADADLDNAKQIVAGADYSGORCTAI 287
Db      231 STEVGKLIMEAANAANKUKVTLELGKSPYIVFDADLDKAVERYIVFGAFGNAQVCIAP 290
Qy      288 KRVLVVEVADLEAEKISENVAKLS-VGDPFNAT--VTPVIDNSAD-FIESLVVDARQ 343
Db      291 SRLVHESIYDEFEVEKTKERVKKLIGDPLDSDTNVYGPLISEQGFDRVLMGYIEDGKE 350
Qy      344 KGAK-----ELNFKDGRLLTPGLPHTLTLMKLMEEPPGPIPIIRVKDAEAV 395
Db      351 EGAKVLCGERDSKKEYLGGGYVQPTIFDVTPEMKIMKEELFGVLPITIKFKDIDEAI 410
Qy      396 ATANKSDPGLQSSVFTRD-FOKAFDIANKLEVGTVAHNNKTRGSPD-NPPFLGLK-GSGA 452
Db      411 ELANDTEYGLAAVYFTKDLIARAFVAKKALEAGIVVNDVCVAABEPOLPFGGVKSSGI 470
Qy      453 GVO-GIRYSIEAMTVKSVL 472
Db      471 GREHGKYGLEBYTEIKVTI 491

```

RESULT 24

```

; US-10-172-585-13
; Sequence 13, Application US/10172585
; Publication No. US2003016200A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Cook, William James
; TITLE OF INVENTION: 32142, 21481, 25964, 21686, NOVEL HUMAN DEHYDROGENASE
; FILE REFERENCE: MNI-134
; CURRENT APPLICATION NUMBER: US/10/172, 585
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US/09/634, 955
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/192, 002
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Aldehyde dehydrogenase family domain
; US-10-172-585-13

```

Query Match 32.4%; Score 775.5; DB 14; Length 492;

Best Local Similarity 38.4%; Pred. No. 1.7e-53;

Matches 192; Conservative 92; Mismatches 167; Indels 49; Gaps 16;

QY 12 EKKSVN--QIILSIDDS--LGFVPAMTREEDVHAMKAGREAL---PAMAAITVERA 65
 DB 1 EWVDSASGKTFEYVNVANKGEVIGRVPEATAEDVDAVAAKAEAFKSGPMMAKVPASERA 60
 QY 66 QYLHKAADIIEERDEKEIATVLAKEISK--AYNASTEYVTRADLITYAA-----EKG 115
 DB 61 RLIRKLAADIIEERDEKEIATVLAKEISKAGTEVERAIDEIRYVGMARKMGERR 120
 QY 116 I--RLSTSDGEGKMDASTGHLKLAIVRQPVGIVLAIPYNPVNLSGSKIPALIGANV 173
 DB 121 VPLSLATDGD-----ELNTRYRREPLGVGVGISPMNPFLLALMKLAPALAAAGT 170
 QY 174 VNFKEPTGSGVSGVLAKAPAEAG--LPAGVFNTITGSGSEIDGYIVHEEVPNFITFG 230
 DB 171 VVLKSEQTPPLTALLAEILIEEGANNLPKGVNVNVPFGAEVQALLSHPIDIKISFTG 230
 QY 231 SPPVQGRIGKLA---GMRPIMLELGKMDAGIYLAADLDNNAKOIYAGAYVDSGQRTAI 287
 DB 231 STEVGLIMEAAAKVLLKVTLELGKSPVIVFDDLDKAVERYVFGAFNAGQVCIAF 290
 QY 288 KEVLVVEEVADELAEKISENVAKL-S-VGDPFDNAT--VTPVIDNSADPTESLVVDAROK 344
 DB 291 SKLIVHESIYDFEVEGLKRVKRLKIGPLDSDTNIVGFLISEQOPDKVLSIEDGKE 350
 QY 345 GAK-----ELNEFKRDGRLLTPGLFDHVTLDMLAEPPGPIIPITRVDAEAVA 396
 DB 351 GAKVLCGGERDESKEVIGGGYVQPTIFDVTPEMKIMEEIFGPVLIIFKFDLDEALE 410
 QY 397 IANKSDPGLQSSVFTPD-FOKAFDINKLEVGTVHINNKTGRPD-NFPPLGL-KSGSG 453
 DB 411 LANDTEYGLAAVYFTKDILARAFVRAKALEAGIVWVDVCVAHAEPQLPFGVHSGSIG 470
 QY 454 VO-GIRYSIEAMTNVKSIVL 472
 DB 471 REHGKGYGLEYTEIKTVI 490

RESULT 25
 US-10-369-493-23067
 ; Sequence 23067, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 23067
 ; LENGTH: 488
 ; TYPE: PRT
 ; ORGANISM: Bacillus subtilis
 ; US-10-369-493-23067

Query Match 31.9%; Score 763; DB 15; Length 488;
 Best local similarity 36.9%; Pred. No. 1,7e-52;
 Matches 170; Conservative 94; Mismatches 179; Indels 18; Gaps 6;
 QY 5 YQNVNGEW--KSNVQIILSIDP-DSSLGFPVAMTREEDVHAMKAGREALPAMAAITV 61
 DB 10 YNPIFGEVWKSQSGMNVKVPADVDVIGVQNSTADVERAVYAAAEAKTAMKRLNG 69
 QY 62 YRAOVYLAADIIEERDEKEIATVLAKEISKAYNASVTEVTRADLIRAAEGRILSTS 121
 DB 70 AERGQVLYTADIMEORLEIIAACATREMGKTLPEAKGETAIGALIRYVAGEGMKLT-- 127

QY 122 ADEGKMDASTGHLKLAIVRQPVGIVLAIPYNPVNLSGSKIPALIGANVMEKPPQ 181
 DB 128 ---GDIVISTDQDALMFTTRVPLGVGVGISPMNPFVLAIPMKMAPALYTGNTVIKPAE 184
 QY 182 GSVGLVLAKAPAEAGLPAGVPNTITGSGSEIDGYIVHEEVPNFITFGSPVQGRIGK 241
 DB 185 TAVTCAKIIACEEESGLPAGVINLVTPGSSVVGQGLAHSDGNATFTGSSQOVGKIIOA 244
 QY 242 AGMR--PIMLELGKMDAGIYLAADLDNNAKOIYAGAYVDSGQRTALKRVLVVEVADE 299
 DB 245 ALARAKTQLEWGNKPIVADDDLEANAHEVITGARRSTGQCTATSRVIVSGIYER 304
 QY 300 LAEKISENVAKL-SVGDPE-DNATVTPVIDNSADPTESLVVDAROKAEL-----NE 351
 DB 305 FKEKLTQRTKQITIDSLKEEDVMWGPISAKQDLWCLSYIEKGQEGASLLIGKEKENG 364
 QY 352 FKRDGRLLTPGLFDHVTLDMLAEPPGPIIPITRVDAEAVAIANKSDPGLQSSVFT 411
 DB 355 KYQNGYVQPAIFDVTSEMTIAQBEIRGPVIALIKVDSIEBALNIANDVKGLSASIFT 424
 QY 412 RFOKAFDINKLEVGTVHINNKTGRGPDNFPPLGLKSGA 452
 DB 425 ENIGMSPFIDEIDAGLVINAESAGVELQAPFGSKQSS 465

RESULT 26
 US-10-369-493-1073
 ; Sequence 1073, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 1073
 ; LENGTH: 463
 ; TYPE: PRT
 ; ORGANISM: Methanococcus jannaschii
 ; US-10-369-493-1073

Query Match 31.7%; Score 759.5; DB 15; Length 463;
 Best local similarity 37.1%; Pred. No. 3.1e-52;
 Matches 175; Conservative 95; Mismatches 183; Indels 19; Gaps 8;
 QY 8 YNNGWKSSVNO--TEILSPIDDSLGFVPAMTREEDVHAMKAGREALPAMAAITVERA 65
 DB 2 FIDGKW---INREDVDVNPYSLVEIKKIPALSTREAKEAIDTAEKYKEMKNLPIRTY 58
 QY 66 QYLHKAADIIEERDEKEIATVLAKEISKAYNASVTEVTRADLIRAAEGRILSTSADG 125
 DB 59 NIIANIAQIKERKEKEKILAIIDAGKPIKQARVEVERISGTFKLAA---FYVKEHREY 115
 QY 126 GKMDASTGHLKLAIVRQPVGIVLAIPYNPVNLSGSKIPALIGANVMEKPPQSGVS 185
 DB 116 IPSD-----DLITFTREVEVGIVGALTTPNFPNLNSAHKIIAIPALANGVIVHHPSSKALV 171
 QY 186 GLVIAK---APAEAGLPAGVFNTITGSGSEIDGYIVHEEVPNFITFGSPVQGRIGK 241
 DB 172 CIELAKIENALKKNVNPVLGVNLLTGAGEVVGDIEIVNEKNMISFTGSSKVGELITRK 231
 QY 242 AGMRPIMLELGKMDAGIYLAADLDNNAKOIYAGAYVDSGQRTALKRVLVVEEVADELA 301
 DB 232 AGFKTIALELGQVNPVIVLKADINKAVNALIKSGFIYAGQVCSVGMILVDESIADKFI 291


```
/ Sequence 228, Application US/10781014
/ Publication No.: US20040180408A1
/ GENERAL INFORMATION:
/ APPLICANT: Pompejus, Markus
/ APPLICANT: Kroeger, Burkhard
/ APPLICANT: Schroder, Hartwig
/ APPLICANT: Zelder, Oskar
/ APPLICANT: Haberer, Gregor
/ TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
/ TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
/ TITLE OF INVENTION: PRODUCTION
/ FILE REFERENCE: BGI-126PCN
/ CURRENT APPLICATION NUMBER: US/10/781,014
/ CURRENT FILING DATE: 2004-02-17
/ PRIOR APPLICATION NUMBER: US 09/602,740
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/141,031
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: 60/143,208
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: 60/151,572
/ PRIOR FILING DATE: 1999-08-31
/ PRIOR APPLICATION NUMBER: DE 19931412.8
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931413.6
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931419.5
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931420.9
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931424.1
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931428.4
/ PRIOR FILING DATE: 1999-07-08
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 784
/ SEQ ID NO 228
/ LENGTH: 484
/ TYPE: PRF
/ ORGANISM: Corynebacterium glutamicum
US-10-781-014-228

Query Match          31.0%; Score 741.5; DB 16; Length 484;
Best Local Similarity 36.9%; Pred. No. 9.2e-51;
Matches 174; Conservative 93; Mismatches 184; Indels 21; Gaps 6;

QY 6 QYVNVSEW-----KSNVQIETLSPIIDSSLGFPVPMTRREVDHAKGREALPMAAL 59
DB 12 KHLGGQWEGNSDRISTN---INPYDSVIAESKQASIDAVDAVEAKKAQAEWMAT 67
QY 60 TVYERAOVYHKAADIIRDKKEIATYLAKEISKAVNASVTEVVRTADLIRYAEERGIRLS 119
DB 68 PAERKRAIITRYRAHELEERREBELVEWLIESGSTSKANLEITLAGNITKESHPFGRV- 126
QY 120 TSADEGKMDASTGHLKAVIRROPVGIYLAIPYNPVNLSGSKIAPALIGNVVMFKP 179
DB 127 ----HGRISPSTPGKENRYVYAKGVGVISPMNPLMSTRSVAPALAVGNAVITKPA 182
QY 180 TGGSVSGVL-AKAPAEAGLPAGVFNITIGRSGEIDYIYHEBEVNFNFTGSTPVGQRI 238
DB 183 SOTPVGVVIPARIFEBAGVPAGVISITVAGAGSEIDHPTVAVPKLISFTSGTPGGRV 242
QY 239 GKLA-----GMRPIMLELGKDGAGIVLADADLDNAKOIYVAGVDSGQRCCTAIKRYLYVE 294
DB 243 GELAINGGPMKYVALLELGNAPFVVLADADIDAAQAAVAGAFHGGQICMSINRYIVDA 302
QY 295 EVADELAERKISENVAKLSVGDPFDNAT-VTPVIDNSADFIETSLVVDAROKAKELNEFK 353
DB 303 AVHDEFLEKFEVEAVKNIPTGDPFAGBGLVGPVINDSQLSGLAKKELAKKKGATVQVEGR 362
QY 354 RGRLLTLPGLFDHYTLDMKLAEEPPGPIIPIIRYVDAEBAVALANKSPFGLOSSFYTFD 413
DB 363 IGRRLVHPVFSVDVTSDMETIAREBEIFGPILSVLKADDEAHAAIETANASDFGLSAAVMSKD 422
```

```
QY 414 FOKAFDIANKLEVGTVHINNKTGRGPDNPFPLGKSGAGVQGIYRSIEAMT 465
DB 423 IDRAAQFALQIDSGVNHINDLTVNDEPRVHMGSKNSGLGFNDMAIEEFT 474

RESULT 30
US-10-369-493-17
/ Sequence 17, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 17
/ LENGTH: 476
/ TYPE: PRF
/ ORGANISM: Aquifex aeolicus
US-10-369-493-17
```

```
Query Match          30.8%; Score 737; DB 15; Length 476;
Best Local Similarity 35.0%; Pred. No. 2.1e-50;
Matches 168; Conservative 96; Mismatches 200; Indels 16; Gaps 6;

QY 4 EYQVNVSEWKSUNVQIETLSPIIDSSLGFPVPMTRREVDHAKGREALPMAALTYE 63
DB 3 EKKMLIGSEWVDEKKEIEVIYPTKRPGRVPGDEKVEKAIERAKGFEIETSLAYE 62
QY 64 RAQVYHKAADIIRDKKEIATYLAKEISKAVNASVTEVVRTADLIRYAEERGIRLS---- 119
DB 63 RYELMRANQOLKEAEFPATLVLEVGKTRREARTEVQRAIQTLIFSAEAKVNGETF 122
QY 120 -TSADEGKMDASTGHLKAVIRROPVGIYLAIPYNPVNLSGSKIAPALIGNVVMFKP 178
DB 123 PIDAPNKK-----GKLGFIYRVGVISAITPFPNPLMSMKH/VADALAAGNAVILKP 176
QY 179 PFGSVSGVLAKAPAEAGLPAGVFNITIGRSGEIDYIYHEBEVNFNFTGSTPVGQRI 238
DB 177 SERTLTPLMGEILLEAGVPPKALSVIRPGG-DVGKAMTHPDVRVVSFTGSRKRVGDI 235
QY 239 GKLAGMRPIMLELGKDGAGIVLADADLDNAKOIYVAGVDSGQRCCTAIKRYLYVEEYAD 298
DB 236 ARQVGIKKVLLELGSNSAIIILHKGDNLEKAEKVYGGIYALAGVYCISVQRFVHEDIYD 295
QY 299 ELAEKISENVAKLSVGDPFDNAT-VTPVIDNSADFIETSLVVDAROKGAK-ELNEFK--R 354
DB 296 ETRKLVKERSVSLKYGDPMNEDTDLGPMIARSEIERIGEWISAVQKAKLEAGGLRCAB 355
QY 355 DGRLLTLPGLFDHYTLDMKLAEEPPGPIIPIIRYVDAEBAVALANKSPFGLOSSFYTFD 414
DB 356 DETTSPFIVSLVPPDSKLFREEARFAPVAVANPNYDIEEALRWNSSSYGQVGFINDV 415
QY 415 OKAFPIANKLEAGTVHINNKTGRGPDNPFPLGKSGAGVQGIYRSIETAMTNVKSIVDM 474
DB 416 KLAACIKETIEAGVVLINBGLPFRADHMPYGVKYSIGREGPKALIEDYIEIKTVIFDL 475

RESULT 31
US-10-369-493-17249
/ Sequence 17249, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
```



```

; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56650
; LENGTH: 482
; TYPE: PR
; ORGANISM: Escherichia coli
US-10-282-122A-56650

Query Match
Best Local Similarity 34.3%; Score 726; DB 15; Length 482;
Matches 164; Conservative 92; Mismatches 206; Indels 16; Gaps 7;

QY 6 QNYVNGEKSSVY--QIEILSPIDSSLGFPVPMTRFEEDHAMKAGREALPMAALTVE 63
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 12 QALINGEWLDANGEMIDVTNPANGDKGVSFPMGDEFRALIDANRLPLPMRALITAKE 71
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 64 RAQYLHKAADIERDKEEIAVLAKEISAVNASVTEVVRTADLIRYAAEGIRLSTSD 123
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 72 RATILRMFNLMEHODDLARLMTLEQKPLAEAKGEISYAASFIEWFAEGKRIYDIT 131
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 124 BEGKXNDASTGHTLAVIRPQVIGVLAIAFYNYPNVLSGSKIAPALIGNVWFKPTQGS 183
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 132 PEGHAD----KRLVITKOPIGVTAITPMNFPMAITRKAGPALAAGCTWLPKASQTP 186
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 184 VSGLVLAFAFAAGLPAGVFNTITGRGSEIGDIYVHEEWNFINFPGVPGORIGKLAG 243
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 187 FSNALALAEIARAQVPAQVFNVTGSAGVNEHTSNPLVKLSFGSTSEIRQLMEQCA 246
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 244 --MRPIMLELGGKADGIVLADLDNAAKQIVAGAYDSGORCTAIKRVLVVEEVADELA 301
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 247 KQIKKVSLELGGNAPFIVPDADLDKAVEGALASKFRNAGQTCVCANRLYVODGYDRPA 306
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 302 EKISNVNALSVDGPPDN--ATVTPVIDNSADFISELVVDAROKAKEL---NEPRDR 357
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 307 EKLQQAIVSKLHTGDGIDNVTIGPLIDERAIVAKVEEHADALEKGRVVGKKAHERGN 366
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 358 ILTPGLFDHTLDMKLAMEEPPGPIPIIRVDAEBAVAIANKSDPGLOSSVFTDFOKA 417
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 367 FPGPIIVDVPANAKVSKETGFLAPLFRPDEADVIAQANDTEGLAAYFYARDLSV 426
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 418 FDIANKLEVGTVHINNKTG--RGPDPNPFGLKSGGAGVQIRYSIEAMTNVKSIVLDM 474
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 427 FVVGELLEYGIVGIN--TGIIISNEVAFPGIGIKASGLGREGSKYGIEDYLEIKYMCIGL 482
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

RESULT 36
US-10-282-122A-46932
; Sequence 46932, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Fdreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848

```

```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46932
; LENGTH: 483
; TYPE: PR
; ORGANISM: Bacillus anthracis
US-10-282-122A-46932

Query Match
Best Local Similarity 35.7%; Score 721.5; DB 15; Length 483;
Matches 171; Conservative 82; Mismatches 203; Indels 23; Gaps 8;

QY 8 YNGWKKSSVNDIELSPIDSSLGFPVPMTRFEEDHAMKAGREALPMAALTVERAQ 67
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 16 YINGEMITLQEQIEVNNPPTKEIFATVPKGVTEAKQVADAHEAFKSKLTADRAK 75
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 68 LHKADIERDKEEIAVLAKEISAVNASVTEVVRTADLIRYAAEGIRLSTADEGK 127
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 76 LKQWFTLIDENKEEIAVLAITHEQKRPFAALGEVYVANSFVWYAEGRV-----YGE 130
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 128 MDASTGHTLAVIRPQVIGVLAIAFYNYPNVLSGSKIAPALIGNVWFKPTQGSVGL 187
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 131 IPASHPNRIILVMKQPVGVMAITPMNFPMAITRKVAPALAGCTAVVKPASQTPPLTXL 190
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 188 VLAKFAEAGLPAGVFNTITGRGSEIGDIYVHEEWNFINFPGVPGORIGKLAG--MR 245
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 191 KLAELAHNADIPKGINIVTGSAKAIADTWMEDEGVRKVSFPGSTSEIGELMASAQTK 250
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 246 PIMLELGGKADGIVLADLDNAAKQIVAGAYDSGORCTAIKRVLVVEEVADELAEKIS 305
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 251 KVSLELGGHAPFIVMDDADLDKAVEAVIGSKFRNAGQTCITNRFVQDEYEALEVEKRE 310
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 306 ENVAKLSVGDPP--DNATYTPVIDNSADFISELVVDAROKAKELNEFRK---DGRLLT 360
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 311 KAVGQLKTKGDDRGDGTGGLIDENAVSKVQEHIEDAIQKGGTVLVGGQKVAELDGHFWQ 370
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 361 P---GLFPHVTILDMKLAMEEPPGPIPIIRVDAEBAVAIANKSPFGLLOSSVFTDFOKA 417
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 371 FTVIGLADMTLCNN---EETGPPVAPVAKFETVEVERANRHTYGLAAYIFTDIOISQ 427
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 418 FDIANKLEVGTVHINNKTG--RGPDPNPFGLKSGGAGVQIRYSIEAMTNVKSIVLDM 474
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 428 FOISEALEYGIIGLNDGL--PSVAQAPFGGKESGIGREGGHFGLIEYLEIKYISLGL 483
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

RESULT 37
US-10-282-122A-60838
; Sequence 60838, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari

```



```
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 18593
/ LENGTH: 480
/ TYPE: PRF
/ ORGANISM: Halobacterium sp. NRC-1
US-10-369-493-18593

Query Match          29.8%; Score 713.5; DB 15; Length 480;
Best Local Similarity 35.0%; Pred. No. 1.6e-48;
Matches 165; Conservative 86; Mismatches 210; Indels 11; Gaps 6;

QY 8 YNNGEKKSVNQIEILSPIDDSLSGFVPAMTREEVNDHAKMGREALPAMALTYVERAGY 67
DB 13 FVAGEVWSGDEPLVPSDLADGGTFASVAADTSDAERALSATGVAHDISETTVERVEM 72
QY 68 LHKADIIRDKKEIATVLAKESKAYNASVTEVVTADLIRYAEGRILSTSADEGK 127
DB 73 LSGINDGIRREDELALEVIREAGKPISSARGEVQSAERFPAVGE-LRLHGTERTG- 130
QY 128 WDASTGK--LAVIRROPVGIYLAIPYVPVNLSSKIAIPALIGNVVFKPPTQGSVS 185
DB 131 --TTAGHEHMOAIVKHEPMGTVCITPVYVPLSTMALQVAPALAAAGNAVIRKPAKTPIS 188
QY 186 GIVLAKAFAPAGIPAGVFTITGRGSEIGDIYVEHEVNFINTGSTPVGQRIKLAGMR 245
DB 189 GAILADIADAGIPDAVYFVPESSVIDDPLASDARVALAINTGSSGAGEHVARQSGIT 248
QY 246 PIMLELGKADGIVLADADLDNNAKQIVAGAYDSGORTCAIKRVLVVEVADLEAKIS 305
DB 249 RLHMEIGSAPAVIPEDDADDAADAATGSLKTAGORCSAVSRVLAHESYHDELVSRI 308
QY 306 ENVAKLSVGDPPD-NATVTPVIDNSADFIESLVVDAROKAKEINEFR--DG-RLIT 360
DB 309 DAAEESIGDLPTDTTGLPLVSADQADWVAELVDDAVRGATVVRGGRHVEDGVHYYE 368
QY 361 PGLFPHVTLDMKLAMEEPPGPIIPIIRVADAEBAVAIAKSDPGLQSSVFTDFOKAFPI 420
DB 369 PTLVADVPDARIIVBEOGFVCAVTVTDEDDAARTANGSELALDAAVFTADHDAMRV 428
QY 421 ANKLVEGTGTHINNKTRGPDNFPFLGLKSGAGVQGIIRYSIEAMTNVKSIVL 472
DB 423 ABRVNAAGAVRINGAPBSHGIGDVPFGGNDASGIGREGIDSTIHFVAREKSIILL 480

RESULT 40
US-10-282-122A-59544
/ Publication 59544, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
```

```
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO 59544
/ LENGTH: 482
/ TYPE: PRF
/ ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59544

Query Match          29.7%; Score 711; DB 15; Length 482;
Best Local Similarity 34.1%; Pred. No. 2.6e-48;
Matches 163; Conservative 90; Mismatches 209; Indels 16; Gaps 7;

QY 6 QNYVNGEKKSVN-QIEILSPIDDSLSGFVPAMTREEVNDHAKMGREALPAMALTYVE 63
DB 12 QALINGQWDAVNGDVIAVNTNPANGEOIGSVKMGADETREAIKANRALPAMRALTAKE 71
QY 64 RAOYHKAADIIRDKKEIATVLAKESKAYNASVTEVVTADLIRYAEGRILSTASD 123
DB 72 RANILIRWFEDLMENQDILARLMTLEOGKPLAEAGEISYASFTIEWFAEGRKIYDGTI 131
QY 124 EGGKMDASTGHTLAVIRROPVGIYLAIPYVPVNLSSKIAIPALIGNVVFKPPTQGS 183
DB 132 PCHQND-----KRLVIRQIPISVTALITPWNPPAMITRKGPALAGCTWVLRKASQRP 186
QY 184 VSGVLAKAFAPAGIPAGVFTITGRGSEIGDIYVEHEVNFINTGSTPVGQRIKLAG 243
DB 187 FSLALAEIANAGIPAGVFNVTGSAGVGGELTSPVRLSTFGSTEIGRLMEQCA 246
QY 244 --MRPIMLELGKADGIVLADADLDNNAKQIVAGAYDSGORTCAIKRVLVVEVADLE 301
DB 247 KDIKVSLELGNAPVIFDDADLDKAVEGALASFRNAGQTCVCANRLYVDGVYDRPA 306
QY 302 EKISENVAKLSVGDPF-DNATVTPVIDNSADFIESLVVDAROKAKEINEFR--DGR 357
DB 307 EKLQAVETKRIQDGLQDGYTTGPIIDEKAVAKVEBHADALAKAKAVVTGKPHALGNN 366
QY 368 ILTPGLFPHVTLDMKLAMEEPPGPIIPIIRVADAEBAVAIAKSDPGLQSSVFTDFOKA 417
DB 367 FPGPITLVVPSAKAKEETFGPLAPLFRFXDEADVIAQANDTFGLAAYFARDLSRV 426
QY 418 PDIANKLEVGTGTHINNKTRGPDNFPFLGLKSGAGVQGIIRYSIEAMTNVKSIVL 474
DB 427 FRVGEALVEYIIGIN--TGIISTEVAPFGVNASGLGRBGRKGIEDYIEIKYMCIGL 482

RESULT 41
US-10-369-493-18711
```

```

; Sequence 18711, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18711
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Halobacterium sp. NRC-1
US-10-369-493-18711

```

```

Query Match      29.7%; Score 710.5; DB 15; Length 489;
Best Local Similarity 35.9%; Pred. No. 2.9e-48;
Matches 167; Conservative 84; Mismatches 199; Indels 15; Gaps 6;

```

```

QY 8 YVNGEM--KSNVQCEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALITVERA 65
DB 15 YIDGIVVPRGDRGLEVDPPSTRELTEIPACTTDDVAAYRAADACREMAADTPQERA 74
QY 66 OYLHKAADIIRDEKEEIAITVLAKEISKAYNASVTEVRTADLIRYAAEGIRLSTADEG 125
DB 75 RIVAAASEILASHREVDVLLATESGTYTKGAASEFASVGTDEAASPTSMN-----G 129
QY 126 GMDASTGHKLAIVIRQPVGIVLAPVNYVNLGSKTAPALIGNVMPKPPQGSYS 185
DB 130 DHKQSVBEKENTIVREPPGVGIIISPMNPLNLSIRAVAPVAAGNSVILKPSASTPT 189
QY 186 -GLVLAFAFAEAGLPAGVFNITGRGSEIGDYIVHEEVPNFTGSPFVGORIGLQAM 244
DB 190 GGLLAAKVFPAAGLBGVFNVTGKSGEIGDRIAGHPADVLSFGOSTAIGVAAITAGE 249
QY 245 RPIM--LEIGKDAIGVLAADLDNAAKQIVAGAYDYSGORCTAIRVLVEEVADELA 302
DB 250 NLALPAMEIGGNVHIVTEDADLDADVDSGVFGSLHSGQICISIRHVHEDVDYVE 309
QY 303 KISENVATLSVDPDPDNATVT--PVLDNNSADPTIESLVDAKQAKELNFRDGLTLP 361
DB 310 RLTERAALPLPTGTADADVTGPIIDESQRDQILYIERTVDAGATLETGDDHGLVVER 369
QY 362 GLFDHVTLDMLKLAEEPPFPIPIIRVKDAEVAIAANKSPGLQSSVFTPRDFOKAFDIA 421
DB 370 TVLSATATMDMAACNEHRRGPVAPVIRPSDDDAVEIAENSTEGVLSGSHAGTGARADIA 429
QY 422 NKLGVGVTHINNKTRGRGPDNFPFLGLKSGAGVQGIIRYSIEAMTN 466
DB 430 DRIDGIMHINDQPVNQEPHLRFGYKNSGLG---RYNGEVVLN 470

```

```

RESULT 42
US-10-369-493-13571
; Sequence 13571, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493

```

```

; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13571
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Thermoplasma volcanium
US-10-369-493-13571

```

```

Query Match      29.5%; Score 705.5; DB 15; Length 473;
Best Local Similarity 34.6%; Pred. No. 6.9e-48;
Matches 167; Conservative 94; Mismatches 194; Indels 27; Gaps 9;

```

```

QY 8 YVNGEMKSSVA--QCEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALITVERA 65
DB 1 YIDGQWVDSNGDITITKPNPVTGEVIGFPAASRDVDVRAEDDAEDAQKMDLQSVRS 60
QY 66 OYLHKAADIIRDEKEEIAITVLAKEISKAYNASVTEVRTADLIRYAAEGIRLSTADEG 125
DB 61 KIIVAKKELIEENRKELENIIIQENGKPVKEAEVDGLDQWQYVAEMARKLNGEVVEG 120
QY 126 GMDASTGHKLAIVIRQPVGIVLAPVNYVNLGSKTAPALIGNVMPKPPQGSYS 183
DB 121 -----TYHRRKIPOYKPYGTVAALTPEWPFAGVARKLAPALITGNTVILKPSDTPGS 175
QY 184 VSLGLAFAFAEAGLPAGVFNITGRGSEIGDYIVHEEVPNFTGSPFVGORIGK--L 241
DB 176 AEWIV--KFIQAGVPKALNVTITGRGSGVIGDYIVHKKVALITMTGSTSGQRIQOAS 233
QY 242 AGMRPIMLELGKDAIGVLAADLDNAAKQIVAGAYDYSGORCTAIRVLVEEVADELA 301
DB 244 ANMAKLMLELGKAPFVWVKDADINNALKTLLMAKYNWAGSCJAABRLVYHEDIYDTFI 293
QY 302 EKISNVAKLSVSPDPDNATVT--PVLDNNSADPTIESLVDAKQAK-----ELNFK 353
DB 294 NKPELSKRIIVGDP--SRADGPIINTLAKTEVEEAESEGARLIYSGSKDLSGPY 352
QY 354 RDGRLTGLGLDPTLDMKLAMEEPPFPIPIIRVKDAEVAIAANKSPGLQSSVFTPRD 413
DB 353 KNGVFLPTIIDNDQSKITQEBIFAVIARKVSDVEEYISLANDSKYGLASTLFTKID 412
QY 414 FQKAFDIANKLEVGVTHINNKTRGRGPDNFP--FLGLKSGAGVQGIIRYSIEAMTNVKSIV 471
DB 413 PNLIFEABERIRFGLYINMP---GPEASQGYHNGFRLTGAQSGSRGILEYLIKNIY 469
QY 472 LD 473
DB 470 VD 471

```

```

RESULT 43
US-09-997-664-3
; Sequence 3, Application US/09997664
; Patent No. US20020151003A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Bassat, Arle
; APPLICANT: Cattermole, Monica
; APPLICANT: Gatenby, Anthony A.
; APPLICANT: Gibson, Katherine J.
; APPLICANT: Ramos-Gonzalez, Isabel
; APPLICANT: Ramos, Juan
; APPLICANT: Sarisastani, Sima
; TITLE OF INVENTION: Method for the Production of p-Hydroxybenzoate in Species of
; FILE REFERENCE: BCI018 US CIP
; CURRENT APPLICATION NUMBER: US/09/997,664
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/585,174
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3

```



```
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 68209
/ LENGTH: 480
/ TYPE: PRF
/ ORGANISM: Pseudomonas putida
US-10-282-122A-68209

Query Match      29.4%; Score 704; DB 15; Length 480;
Best Local Similarity 34.0%; Pred. No. 9,3e-48;
Matches 162; Conservative 91; Mismatches 209; Indels 14; Gaps 7;

QY      6 QNYVNGEMKSSVNV--QIELSPIDSSSGFVPMATREEDVHAMKAGREALPMAALTYVE 63
DB      12 QYVINGEWDADNGOTIKVTNPATGSEVIGTVEPKMGTAETRRALTEADKALPMPRALTYAKE 71
QY      64 RAQYLAHKAADIERDKESIAFVLAKEISKAVNASYTEVVRTADLIRYAAESGIRLSTSDAD 123
DB      72 RSAKLRMEELMIENQDOLARLMTTEQCKPLAEKGEIYVAASFIEWPAEAKRIYGDIT 131
QY      124 ECGKMDASTGKHLAVIRROPGVIVLAIAPYNYVNLGSKIAIPALIGNVVFKPPTQGS 183
DB      132 PEHQPD-----RLIYIKQPIGVTAALTPWNPAMITRKAGPALAAGCTVWLKPSQTP 186
QY      164 VSGVLAKAFAGAGPAGVFNITITGRSGSIGDYIYHEHEVNFINTGSPVQGRIGK--L 241
DB      187 YSALLVELAHRAAGIPAGVLSVATSGAGVGGELTGNLSVRLSLFTGSGTEIGRQIMEBCA 246
QY      242 AGMRPIMLELGGKDGIVLADADLNNAAKOIVAGAYDYGQRCTAIKRYLVVEEVADELA 301
DB      247 KQIKVYSLELGGNAPFIYDDADLDKAVGGAIISTRNNGQTCVCANRIYVDGVIYDARA 306
QY      302 EKISENVAKLSVGDPEFNATVT--PVIDNSADFIESLVVDARQKAKELNEFKR--DGRLL 359
DB      307 EKDLAAVAKLAKIGNGLEBGTITGPIIDGKAVAKVQEHIDAVSKAKVLGSGKLEGNF 366
QY      360 TPGGLDHTYLDKMLAMEBEPGILPIIRKYDAEBAVAIAKNSDPGIGSSVFTPDQKAPD 419
DB      367 EPTLIVDPKTAIVAKKEETFGPLAPLFRFKDAEVLAMSNDEFGIASYFYARDMSRVRR 426
QY      420 IANKLEVGTVHINNKTGR--GPDNFPPLGKSGAGVQIGRYSIEAMTVNKSIVLDM 474
DB      427 VAEALIEGVNGIN--TGLISNEVAPFGIKASGLRGESKGIETYLEIKYICISV 480

RESULT 46
US-10-369-493-17169
/ Sequence 17169, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
```

```
/ SEQ ID NO 17169
/ LENGTH: 468
/ TYPE: PRF
/ ORGANISM: Bacillus halodurans
US-10-369-493-17169

Query Match      29.1%; Score 697.5; DB 15; Length 468;
Best Local Similarity 33.4%; Pred. No. 3e-47;
Matches 158; Conservative 93; Mismatches 209; Indels 13; Gaps 5;

QY      8 YNNGEMKSSVNOIELSPIDSSSGFVPMATREEDVHAMKAGREALPMAALTYERQY 67
DB      4 YINGEMKSSGKTLVDVNPATGSEVIDYVFAKKEKELVSAAYEFPMSSQTAESERSY 63
QY      68 LHKADIERDKESIAFVLAKEISKAVNASYTEVVRTADLIRYAAESGIRLSTSDADCGK 127
DB      64 LMRWFQILDEQODELGEIMTTEQCKPLAEALGEQYANSFIQWIAEAKRI-----YDPT 118
QY      128 MDASTGKHLAVIRROPGVIVLAIAPYNYVNLGSKIAIPALIGNVVFKPPTQGSVGL 187
DB      119 IPASAINRILIVQKQPGVIVAIIPWNPAMITRKAVPALAAGCTALVKDAEQTPAL 178
QY      188 VLAARAFAGAGPAGVFNITITGRSGSIGDYIYHEHEVNFINTGSPVQGRIG--GTLAAMR 245
DB      179 KLAQLAEEAGIPAGVLANITGNADIGRAWLEDSRVRKITPTGTEVGKLMRGAQTVK 238
QY      246 PIMLELGGKDGIVLADADLNNAAKOIVAGAYDYGQRCTAIKRYLVVEEVADELAEKIS 305
DB      239 KISLELGHAPFITMDANLEEAVDQYASFRNAGQTCVCANRIYVAREIAEATEKFA 298
QY      306 ENVAKLSVGDPE--DNATVTPVIDNSADFIESLVVDARQKAKELNEFKRDGRLLTPGLF 364
DB      299 AKVVELKVGNGLEBGTITGPIIDKAAVEKAHIDALKKGGQTVVGGRWNTNHPFEPTII 358
QY      365 DHTYLDKMLAMEBEPGILPIIRKYDAEBAVAIAKNSDPGIGSSVFTPDQKAPIANKL 424
DB      359 TGADEMLCMNEETFGPLAPVATPDTEEEVERANHTPYGAAAYFTENIRAIARLSKL 418
QY      425 EYGVTHINNKTGRGPD--NFPFLGKSGAGVQIGRYSIEAMTVNKSIVLDMK 475
DB      419 EYGVGVNDGM---PSVAQAPFGWKESGLRGEGKGIIEYLEVYKYSVAFK 468

RESULT 47
US-10-369-493-18996
/ Sequence 18996, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 18996
/ LENGTH: 473
/ TYPE: PRF
/ ORGANISM: Anabaena PCC7120
US-10-369-493-18996

Query Match      29.0%; Score 694; DB 15; Length 473;
Best Local Similarity 32.6%; Pred. No. 5,8e-47;
Matches 157; Conservative 102; Mismatches 198; Indels 24; Gaps 8;

QY      8 YNNGEMKSSVNOIELSPIDSSSGFVPMATREEDVHAMKAGREALPMAALTYVER 64
```

Db 1 YINGOMNATEETLINSHPADKEIVATFPRSQADDVDAVAARQAYSGWRKVAPAR 60
Qy 65 AAYLHKADIIERDKEEITVLAKEISKAYNASVEVATADLIRAAEGRLSTSADE 124
Db 61 AAYIRFVGGELLQHEEELAQLSIREKGPDLTERAGDVQGVCAFYSSAGEGRLLF----- 115
Qy 125 GSKMDASTHKLAVIRROPVGIYLAIPYNYPVNLSSGKIAPALIGANVVMFKPPTQGSV 184
Db 116 GQTTSEEMNKRMKRMKMPICVICALITTPNNFVAITCMKAMPALVCGNTVILKPADDTA 175
Qy 185 SGVLAKAPAEAGLPAGVNTITGRGSEIGDYIVEHEVNFNFTGSTPVGORIGLACM 244
Db 176 CATKLEIFFAAGLPFGVINLVHVGEEYKALVEHPNIDLVSTSSATGAYVGETGR 235
Qy 245 --RPMLEIGKDGAGIVLADADLDAKQIVAGANDYSGORCTAIRVILVEBVADELAE 302
Db 236 THKRCLEWGGKNAQVMDADLELDALMGCAFQGTGCTATSRLLIHDIXEKFT 295
Qy 303 KISENVAKLSVGDPPFNAT-VTPVIDNSADPLESI-VVDAROKAKEL--NEFKDGR- 358
Db 236 MUREKTSQRLAGTEPEPTDIGITINNROLRVHEVNYIAREGAKILIGGEIATEGOLQ 355
Qy 359 ---LTPGLFDHVTLDKLAWEPPGPIIPIRVKDAEBAVALANKSDGLQSSVETRD 414
Db 356 QGYFQPTILDVTPQMRVAREEIRFPVALIEVSTFEBAIALINDTKYGLSSVYTRDI 415
Qy 415 QKAFDIANKLEVGVTHINKTGRGPDNPFPLGKSGAGVQIGRYSIEAMTVKSI VLDM 474
Db 416 NNAFVAMRDIEMAGIYINGPITGAIEVHLDPFGVKQKGNHR-----EAGTALDVFTW 469
Qy 475 K 475
Db 470 K 470

RESULT 48

US-10-369-493-8741
; Sequence 8741, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Cao, Yongwei
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052) B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8741
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Ralstonia metallidurans
US-10-369-493-8741

Query Match 29.0%; Score 694; DB 15; Length 477;

Best Local Similarity 35.5%; Pred. No. 5,9e-47; Indels 34; Gaps 10;

Matches 173; Conservative 83; Mismatches 197;

Qy 5 YONYVNGEWSKSVNOIEILSPIDDSLGFVPMTEEEVD-HAMKAGREALPA----- 55

Db 3 YANYIDGECDSASG---RTVDNVA---PADTDIVSRHAASDARDAAAIVAAAAAF 54

Qy 56 --WALTTERAQYLHKADIIERDKEEITVLAKEISCAVNASVEVTRADLIRYAAE 113

Db 55 DQMKTPIGKRAKIINDAAHLEANADTATELTREEGKALVADDEVRSQOTRFYAV 114

Qy 114 EGRILSTSDGSKMDASTGHLAVIRROPVGIYLAIPYNYPVNLSSGKIAPALIGAV 173

Db 115 EGQTSGESGYNDPDD-----MLVYSLEPPLGVTVIISPWNFPVSI PARKIAPALITGNT 169
Qy 174 VMFKPPTQGSVGLVLAAPAEAGLPAGVNTITGRGSEIGDYIYHEHEVNFNFTGSTP 233
Db 170 VVFKPSSAPLSGLYLAEPVAGIPKGYLNFILTSAAEVGTYIESRVRVAVSFTGSTS 229
Qy 224 VQGRIGKLAKMP-IMLEIGKDGAGIVLADADLDAKQIVAGANDYSGORCTAIRVIL 292
Db 230 ABEQIHKSVPMTRRQOMELGKRNPLIEMEDADLDAVDAVAGLSLSQAQCTGSRILV 289
Qy 293 VEEVADLEAKISENVAKLSVGDPPD-NATVTPVIDNSADPLESI-VVDAROK-----GA 346
Db 290 MBDVAQAFTEKLVAAYVXALKIGSGMTAGMDGLPATTRKQLETVLVYIAIGKQEAETLLCGG 349
Qy 347 KELAEPK-RDGRLLTPGLFDHVTLDKLAWEPPGPIIPIRVKDAEBAVALANKSDPGL 405
Db 350 EQLSDGDPAHGYVVAFTVFTDTQSMRIAREEIRFPVALIEVDSYADAIAQANDTEYGL 409
Qy 406 QSSVTRDQKAFDIANKLEVGVTHINKTGRGPDNPFPLGKSGAGV--QGIRYSTEA 463
Db 410 SAAIATRRMRVMDFTRIDIESGTVKINRTTGNLVNAPPGLKRSSTFRBSGAGLEF 469
Qy 464 MTNPKSI 470
Db 470 YTOIKTV 476

RESULT 49

US-10-282-122A-69547
; Sequence 69547, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69547
; LENGTH: 480


```

; FILE REFERENCE: ELITRA_034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75217
; LENGTH: 482
; TYPE: PRF
; ORGANISM: Salmonella typhimurium
US-10-282-122A-75217

```

```

Query Match      28.9%; Score 692; DB 15; Length 482;
Best Local Similarity 33.3%; Pred. No. 8.6e-47;
Matches 159; Conservative 95; Mismatches 208; Indels 16; Gaps 7;

```

```

QY 6 QNYNGEKKSSV--VNOIEILSPIDSSLGFPVPMTEBVDHMKAGREALPMAALTVE 63
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 12 QAFIDGAWWDADNGQTIKNNPATGEIIGTVPKMGAETRRALTEADKALPAMRALTAKE 71
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 64 RAQYLHKADIERDKEIATVLAKEISKAVNASVTEVVRTADLIRYAABEGIRLSTSD 123
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 72 RANILRRWELMTIEMODDLARLMTLEQKPLAEAKGEIYVAASFIEWFAEKKRIYGDVI 131
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 124 EGKMDASTGHLAVIRQPVGIVLAIPYNYPNLSGSKIAPALIGNVNMFKEPTQGS 183
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 132 PGHQAD----KRLVITKQPIGVTAITPWNFPAMITRKAGPALAAGCTWVLKPAQT 186
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 184 VSGVLAKAFAPAGLPAGVFNITITGRGSEIGDIYIVHEEVENFTNGSTIPVQORIGK 243
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 187 FSAIALAEIAQRAQIPAGVFNVTGSGADIGSELTSNPLVRKLSTFGSTIEIGRLMEQCA 246
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 244 --MRPIMLEGGDAGIVLADADLNAKQIVAGAYDSGRCSTAIKRVLYVEEVADELA 301
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 247 KQIKKVSLEGGNAPPIVFDADLDKAVEGALISKRRNAGQTCVCANRLYDQGVYDRFA 306
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 302 EKISENVAKLSVDPDFD-NATVTPVIDNSADFIESLVVDARQKAKEI--NEFRDGR 357
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 307 EKUNQAVNKLAVDGDQADVAIGPLIDEKAVAKVQEHIALEKGRVITGSAHLGCGN 366
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 358 LLTPGLFDHYTLDMKLAMEEPFGIILPIIRKDAEENALANKSDPGLQSSVFTRDPQKA 417
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 367 FPGPTLAVPDPDAKAKAKETFGPLAPLFRFSDADVIRQANTFEGLAAYFARLDISYV 426
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 418 FDLANKLEVGTVIHNKKG-RGPDNPFGLKSGAGVQGRISYEAMTVKSIIVLDM 474
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 427 FRVGEALIEYGVGIN--TGLISNEVAPFGSIKASGLGRBSKYGIEDYLEIKIMCIGL 482
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 52
US-10-369-493-13957
; Sequence 13957, Application US/10369493
; Publication No. US20030233675A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13957
; LENGTH: 478
; TYPE: PRF
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-13957

```

```

Query Match      28.9%; Score 691; DB 15; Length 478;
Best Local Similarity 33.5%; Pred. No. 1e-46;
Matches 159; Conservative 92; Mismatches 209; Indels 14; Gaps 7;

```

```

QY 6 QNYNGEKKSSV--QIEILSPIDSSLGFPVPMTEBVDHMKAGREALPMAALTVE 63
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 11 QAFIDGAWWDADNGQTIKNNPATGEIIGTVPKMGAETRRALTEADKALPAMRALTAKE 70
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 64 RAQYLHKADIERDKEIATVLAKEISKAVNASVTEVVRTADLIRYAABEGIRLSTSD 123
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 71 RANILRRWELMTIEMODDLARLMTLEQKPLAEAKGEIYVAASFIEWFAEKKRIYGDVI 130
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 124 EGKMDASTGHLAVIRQPVGIVLAIPYNYPNLSGSKIAPALIGNVNMFKEPTQGS 183
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 131 PGHQAD----KRLVITKQPIGVTAITPWNFPAMITRKAGPALAAGCTWVLKPAQT 185
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 184 VSGVLAKAFAPAGLPAGVFNITITGRGSEIGDIYIVHEEVENFTNGSTIPVQORIGK-L 241
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 186 FSAIALAEIAQRAQIPAGVFNVTGSGADIGSELTSNPLVRKLSTFGSTIEIGRLMEQCA 245
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 242 AGRPIMLEGGDAGIVLADADLNAKQIVAGAYDSGRCSTAIKRVLYVEEVADELA 301
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 246 KQIKKVSLEGGNAPPIVFDADLDKAVEGALISKRRNAGQTCVCANRLYDQGVYDRFA 305
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 302 EKISENVAKLSVDPDFD-NATVTPVIDNSADFIESLVVDARQKAKEI--NEFRDGR 359
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 306 EKUNQAVNKLAVDGDQADVAIGPLIDEKAVAKVQEHIAVSKATVLAGKPMEGNPF 365
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 360 TGLFDHYTLDMKLAMEEPFGIILPIIRKDAEENALANKSDPGLQSSVFTRDPQKAD 419
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 366 EPTIILNPKDAVAVAKKEETFGPLAPLFRFXDADVIAMSNDEFGIASYFARLDISYV 425
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 420 IANKLEVGTVIHNKKG-RGPDNPFGLKSGAGVQGRISYEAMTVKSIIVL 472
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 426 VABALEYGVGIN--TGLISNEVAPFGSIKASGLGRBSKYGIEDYLEIKIYCL 477
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 53
US-10-282-122A-51224
; Sequence 51224, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.

```

```

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 51224
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-10-282-122A-51224

Query Match      28.8%; Score 690; DB 15; Length 489;
Best Local Similarity 33.4%; Pred. No. 1,3e-46;
Matches 161; Conservative 93; Mismatches 200; Indels 28; Gaps 8;

QY      8 YNNGEKSXSVNOIELSPIDSSLG---FVPAMTSEVDHAKAGREALPMAALTYVE 63
      19 YIDGKRVGANGNST--PYDNPSTGKTITISVPLGSKKEAQAIDAAEALPMSATGXE 76
DB      64 RAQYLHKAADIIERDKEEIAATYLAKEISKAYNASTVEVTRADLIRYAAEGIRLSTSD 123
      77 RAAILLKMAQLMNENQODLAVIMTSEQKVPPEAGEIYAASFLEWPAEAKRI----- 131
QY      124 EGGKMDASTGHTLAIVRROPVGIVALAIPYNYVNLSSKIAPIALIGNVVMFKPTQGS 183
DB      132 DDIIQSPKGGORLWMLKOPIGVTAITWMNFPAAMITRKVGPAALAGCTMVVKPAQQTP 191
QY      184 VEGLVYAKAFEAAGLPAGVFNITTGSGSEIGDYIVHEEYVNFNFTGSTPVGORIKLAG 243
DB      192 LTRALLAVAVLAGVAGVAFQVVTGSSREIGALCESEYVRKLSFSGSEIVGRTLMEOCA 251
QY      244 --MRPIMLELGGKADGIVLADADLDNAAKQIVAGAYDYSQORCTAIKRYLVVEEVADELA 301
DB      252 PTKIKLSTELGNAPIVFDADLDRAIDGILLASKRYNAGQTCVCANRIYVQAGYVEVA 311
QY      302 EKISEVVALISVGDPPD-NATYTPVVIDNSADPISLIVDAROKAKEL---NEFRDOR 357
DB      312 KSLVEVVMAMKVGDDGEGVTOGPLLDTAAVAKVQHIDATNAHGAQVLAGGEPAHAGGS 371
QY      358 LITPGLFDHVTLDMLKAMEEPPGILPIIRVKDAEAVAIANKSDGLOSVVTRPFOKA 417
DB      372 FPQPTVVRVTOGSMRAVEETGFPVAPLTKFSEDEVIGMANNTITGLAAYFYTRDYARI 431
QY      418 FDIANKLEVGTVHIN-----NKTGRGPDNFPFLGLKGSAGVQGIRYSIEAMTNVSIYL 472
DB      432 WRVSEALEYGIIVGINIGIISNEVG-----PFGVGVSQIGRGSGSYGIEDYELKYLAV 485
QY      473 DM 474
DB      486 DL 487

```

```

RESULT 54
US-10-369-493-11184
; Sequence 11184, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11184
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Methanosarcina mazei
US-10-369-493-11184

Query Match      28.7%; Score 687; DB 15; Length 480;
Best Local Similarity 32.1%; Pred. No. 2.2e-46;
Matches 155; Conservative 105; Mismatches 207; Indels 16; Gaps 6;

QY      1 LTKXYONYNGEBWK--SSVNOIELSPIDSSLGFPAMTSEVDHAKAGREALPMAA 58
      3 MVQBYKLFIFGEFKDOSTGETFEDINPATLNLATIQVAGADVDMAVEAAEAGFRWNE 62
DB      59 LTVYRAQYLHKAADIIERDKEEIAATYLAKEISKAYNASTVEVTRADLIRYAAEGIRL 118
      63 VPAARRAVLFPRAAIIQERKEEISVMTKEMGKVLPTRGDVQDAIDITYAAGEGRM 122
QY      119 STSADEGKMDASTGHTLAIVRROPVGIVALAIPYNYVNLSSKIAPIALIGNVVMFKP 178
DB      123 L-----GETTSELKEKCMITLRPIGVGLITPNFPIALPAMKVMALVAGNAIVFKP 177
QY      179 PTQGSVSGIVLAKAFEAAGLPAGVFNITTGSGSEIGDYIVHEEYVNFNFTGSTPVGQRI 238
DB      178 ASDTFLAFKLEIVSEAGLPFGVINLVTGCGTVGKAVVOHPRIKALISFGSLDTGKWI 237
QY      239 GGLAG--MRPIMLELGGKADGIVLADADLDNAAKQIVAGAYDYSQORCTAIKRYLVVEV 296
DB      238 MEECSKTKRVSLTEIGKQNPVIVMDADLELALBGVLMGARFGTGGORCTATSRLLHKEI 297
QY      297 ADELAEKISENVAKLSVGD--FDNATVTPVIDNSADPISLIVDAROKAKELNEFKR- 354
DB      298 KDEFTKRLLAARAKALSTIGGILPETIDIPVINKQLEIKERYVKIKREGATILYGKRI 357
QY      355 ----DGRLLITPGLFDHVTLDMLKAMEEPPGILPIIRVKDAEAVAIANKSDPGLQSSVF 410
DB      358 DPGLVGFYFEPITFDVAPDMRIAGEEIFGVLGIFTVSDLEAATILANSTKYGSATY 417
QY      411 TRDPQKADPIANKLEVEGVVHINKTGRGPDNFPFLGLKGSAGV--QGIRYSIEAMTNVKS 469
DB      418 TGNIGNAARAIEKVEAGITVVAPTIGAIVHLFGVGVGDTGNGFREAGTEAVKEPSEVKA 477
QY      470 IYL 472
DB      478 VTI 480

```

```

RESULT 55
US-10-369-493-8312
; Sequence 8312, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

```

```

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8312
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Thermobifida fusca
; US-10-369-493-8312

Query Match      28.7%; Score 686; DB 15; Length 467;
Best Local Similarity 36.1%; Pred. No. 2,5e-46;
Matches 168; Conservative 91; Mismatches 203; Indels 4; Gaps 4;

QY 8 YVNGEKSVNQIEILSPIDSSIGFVPAMTREVDMKAGREALPAMAAITVYERAOY 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 5 WLKSAFTGDNELTVNPTGCVATVAVPSPEOVQAAAHGALDATALPAHRAVA 64

QY 68 LHKADIIERDKKEIATVLAKETSKAYNASVTEVVTADLIRYAEGRISTSADEGK 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 65 LAHVAARLORSEIEIQLITAESEKPIIMARAEVGAVSFTAAEE-VARDSGRVORLD 123

QY 128 MDASTGHKLAIVITROVGIYLAIPYNPVNLGSGKIAPALIGNVMPFPQVSGSL 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 124 TDSAGERRALIRRVPGAVLALSPFPLNLVAHKAVALIAGAPLVKPADPLSLAL 183

QY 188 VLAKAFEAAGLPAGVFNTITGRGSEIGDYVEHEEVNFINFTGSPVQRIKLAGMRP 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 184 VDEIILAEITDLPRAGASVLP-MENELAQPLITBERLPVISTFG-QFGWELRGMANKRY 241

QY 248 MLELGKDGAGIVLADADLDNAKQIVAGAYDYGQRCTAIKRYLYVEEVADELAEKIS 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 242 TELSGNAALAVLADADLDMARARIALFGNNQAGQVCIGQRYVEDAVDEPAALVQR 301

QY 308 VALTSGDPPDNAT-VTPVIDDNSADPIESLVYDARQKKEINERDORLLTPGLF 366
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 302 VEEVLVTGDPDADEATQVGPLINEAAARVEQWNEAAGATVLTGTGCVTMAPTVLAD 361

QY 367 VTIDMKLAMEEPFGPIIPITRVDAEBAVAIAKSDPGLOSSVFTFDFOKAFDIANKLEY 426
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 362 VPPDAKVCHEVEGFWVWLRVSGVDEAFAPVNNSKTGLDAGVFTDLATAPAHRELKY 421

QY 427 GTVHINKTGRGPDNPFPLGLKSGAGVQGIKSYIEAMTNVKSIVL 472
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 422 GGVITIGDVPYRADQMPYGVGSKSGVKGESVRSAMMDLDERVMVL 467

RESULT 56
; US-10-369-493-4282
; Sequence 4282, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

```

```

; SEQ ID NO 4282
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
; US-10-369-493-4282

Query Match      28.5%; Score 682.5; DB 15; Length 458;
Best Local Similarity 36.1%; Pred. No. 4.7e-46;
Matches 168; Conservative 84; Mismatches 199; Indels 15; Gaps 7;

QY 11 GEMKSVNQIEILSPIDSSIGFVPAMTREVDMKAGREALPAMAAITVYERAOYLHK 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 GERVTRPRLDVDFPTGTVGVPLAS---VDDVRAAREVAAYAOQLTRERSQILHR 57

QY 71 AADIIERDKKEIATVLAKETSKAYNASVTEVVTADLIRYAEGRISTSADEGKMDA 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 58 AASIMKERTEEASDILSIESGLSKDSRYEIRGVADVLFASIEALR-----DDGQSFSC 112

QY 131 SF---GHKLAVI-RROPV-GIVLAIPYNPVNLGSGKIAPALIGNVMPFPQVSGVS 185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 113 DLTPHGKRRVRSQREPLAGVIALTFPHHPNQVAHKAIPALATNNRVLLKPSKVLPS 172

QY 186 GLVAKAFEAAGLPAGVFNTITGRGSEIGDYVEHEEVNFINFTGSPVQRIKLAGMR 245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 173 ALYIADVLYEAGLPAPMLQVLTGDPREIADDELITPHLAEVFTFGVAIGKTYIAAPAAVR 232

QY 246 PMLELGKDGAGIVLADADLDNAKQIVAGAYDYGQRCTAIKRYLYVEEVADELAEKIS 305
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 233 RVLLEIGDNDPLIVLDDADLERRAATLAVQSGYKNSGQRTAVKRLVQKSVADFTDLVV 292

QY 306 ENVAKLSVGDPPD-NATVTPVIDDNSADPIESLVYDARQKKEINERDORLLTPGLF 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 293 EXTRAWTFEDPPDASQWGTVIDVAAALQFEARVNEAASGARLLTGNNRGAALYAPVL 352

QY 365 DHVTLDMKLAAMEEPFGPIIPITRVDAEBAVAIAKSDPGLOSSVFTFDFOKAFDIANKL 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 353 DEVDSBMLTVREBTGCPVPIITPTDLDAIINSGTAIGSSGVCOTNRQDAITRINEL 412

QY 425 EVGTVHINKTGRGPDNPFPLGLKSGAGV-QGIKSYIEAMTNVKS 469
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 413 RVGTVMVEVPGYRIELTFPGGIDKSGLGKGVQDEAMKSPFNLT 458

RESULT 57
; US-10-369-493-7037
; Sequence 7037, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7037
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
; US-10-369-493-7037

Query Match      28.5%; Score 682.5; DB 15; Length 458;
Best Local Similarity 36.1%; Pred. No. 4.7e-46;
Matches 168; Conservative 84; Mismatches 199; Indels 15; Gaps 7;

QY 11 GEMKSVNQIEILSPIDSSIGFVPAMTREVDMKAGREALPAMAAITVYERAOYLHK 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



```

Db      132 PGHQT-----KRLVTKQPIGVTAATPWNPPSAMITRKAGPALAAGCTMVLKPSQTP 186
QY      184 VSGLVLAFAAEGLPAGVNTTGRGSEIGDIYVEHEVNFNFTGSPVGRIGKLAG 243
Db      187 FSLALAEALQRAAGIAGVNVVTSAGDIGELTSPVLRKLSFTGSTIEIGRLMEQCA 246
QY      244 --NRPIMLELGGKAGIADLADLDNAKOIVAGAVDSGRCATIKRVLYVEEVADELA 301
Db      247 KDIKVVSLGAGNAPFIVPDDADLDKAVGALASKFRNAGQTCVCANRLYVDGVDRFA 306
QY      302 EKISENVAKLSVGDPPD-NATVTPVIDNSADFIESLVVDARQKAKEL---NEFRDGR 357
Db      307 EKLNQAVNKLAVDGLQADVAIGPLIDEKAVAKVGEHIALDEKAGRVITGGEAHLKGN 366
QY      358 LITPGLEFDTLDMKLAMEEPFGILPIIRVKAEEAVAIANKSDPGLQSSVTRDFQKA 417
Db      367 FPOPTIADVPDNAAKAKETFGPLAPLFRFSEADVIRQANDTEGLAAYFYARDLSRV 426
QY      418 FDIANKLEVGTVHINKTG-RGPDNPFPLGLKSGAGVQIGIRYSIEAMTVKSIYLD 474
Db      427 FRVGEALFYGIVGIN-TGIISNEVAPFGGIKASGLGREGSKYGIEDYLEIKYMCIGL 482

```

RESULT 60

US-10-282-122A-75920

```

/ Sequence 75920, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forey, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282.122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 75920
/ LENGTH: 482
/ TYPE: PRT
/ ORGANISM: Salmonella typhi
US-10-282-122A-75920

```

```

Query Match      28.5%; Score 682; DB 15; Length 482;
Best Local Similarity 33.1%; Pred. No. 5,5e-46;
Matches 188; Conservative 95; Mismatches 209; Indels 16; Gaps 7;

QY      6 QNVNGEMKSS--VNOIEILSPIDSSLGFPVAMTRREEVDHAMKAGREALPMAALTYVE 63
Db      12 QAFIDQDMADAGGVITPVSNPANGPKLGNVFRMGAEFRDAIIMANRALPMAALTAKE 71
QY      64 RAOYLHKAADIIRDEKEEATVLAKEISKAYNASTVEVTRIDLIRYAAEEGRILSTSD 123
Db      72 RANILRRFNMIMMEHODDLARLMTLEQGRPLAEAKGEISYAASFIEWFAEEGRIRYDGT 131
QY      124 EGGKMDASGHGLAVRRQPVGIVLAAPYNPVNLSSGIAPALIGVWVMEKPTQGS 183
Db      132 PGHQT-----KRLVTKQPIGVTAATPWNPPSAMITRKAGPALAAGCTMVLKPSQTP 186
QY      184 VSGLVLAFAAEGLPAGVNTTGRGSEIGDIYVEHEVNFNFTGSPVGRIGKLAG 243
Db      187 FSLALAEALQRAAGIAGVNVVTSAGDIGELTSPVLRKLSFTGSTIEIGRLMEQCA 246
QY      244 --NRPIMLELGGKAGIADLADLDNAKOIVAGAVDSGRCATIKRVLYVEEVADELA 301
Db      247 KDIKVVSLGAGNAPFIVPDDADLDKAVGALASKFRNAGQTCVCANRLYVDGVDRFA 306
QY      302 EKISENVAKLSVGDPPD-NATVTPVIDNSADFIESLVVDARQKAKEL---NEFRDGR 357
Db      307 EKLNQAVNKLAVDGLQADVAIGPLIDEKAVAKVGEHIALDEKAGRVITGGEAHLKGN 366
QY      358 LITPGLEFDTLDMKLAMEEPFGILPIIRVKAEEAVAIANKSDPGLQSSVTRDFQKA 417
Db      367 FPOPTIADVPDNAAKAKETFGPLAPLFRFSEADVIRQANDTEGLAAYFYARDLSRV 426
QY      418 FDIANKLEVGTVHINKTG-RGPDNPFPLGLKSGAGVQIGIRYSIEAMTVKSIYLD 474
Db      427 FRVGEALFYGIVGIN-TGIISNEVAPFGGIKASGLGREGSKYGIEDYLEIKYMCIGL 482

```

RESULT 61

US-10-369-493-16681

```

/ Sequence 16681, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 16681
/ LENGTH: 460
/ TYPE: PRT
/ ORGANISM: Bacillus thuringiensis
US-10-369-493-16681

```

```

Query Match      28.4%; Score 680.5; DB 15; Length 460;
Best Local Similarity 34.6%; Pred. No. 6,8e-46;
Matches 161; Conservative 82; Mismatches 205; Indels 17; Gaps 6;

QY      19 QTEILSPIDSSLGFPVAMTRREEVDHAMKAGREALPMAALTYVEVRAOYLHKAADIIR 78
Db      4 QIEVNNPATKEIFATVFKGCVTEBAKQAVDAAEAFRTMSKLTAAADRATLTKKFMPLIDQ 63
QY      79 KEIATVLAKEISKAYNASTVEVTRADLIRYAAEEGRILSTSADEGGRMDASTGHKLAV 138
Db      64 KEIATVLAKEISKAYNASTVEVTRADLIRYAAEEGRILSTSADEGGRMDASTGHKLAV 138

```



```

; LENGTH: 474
; TYPE: PRT
; ORGANISM: Rhodococcus ruber
US-10-273-051-12

Query Match      28.1%; Score 672; DB 14; Length 474;
Best Local Similarity 35.5%; Pred. No. 3,4e-45;
Matches 169; Conservative 84; Mismatches 197; Indels 26; Gaps 9

OY 8 YNNGEMKSVNO--TEIISPIDSSIGFVPAMTREBEVDHAKRAGREALPAMAALTYVERA 65
Db 10 YINGSWVASTSKTIVLEINPATBEVIGTVPDGTADVDAAVAARAAPFGWASTPDKXA 69
OY 66 QYLRK-AADIIERDKBEIATVTLAKETSKAYNASTVEVETADLIRYAEEGRILSTADE 124
Db 70 QYLRAAAGIADRSEDLARTTISAEMGAPLSFQAOMQVPLPINSFSHAAVAESFPFERT 129
OY 125 GKKMDASGKHLAVIRROPVGIVLAPRYNPNVLSGSLIARALIGNNVMPKPTQGSV 184
Db 130 GS-----SVTVEREPIGVGAITPMNYPPLHQIAAVALAAGNTIVKPSVADL 179
OY 185 SGLVLAKAPAEAGLDPAGVFNITITGSGSEIGDYIVHEEYVNFINTGSPVGORIGLAG- 243
Db 180 NAMMELIIDAAGVAPAGVFNVLVSGRPVVGELASHHEVDMLSTFQSTNAKGRVSLAAQ 239
OY 244 -MRPLMELGCKDAGIVTADADLDLNAKOIVAGAYDYSQQRCTAIRVILVEEVADELAE 302
Db 240 TVKRVALEIGKGSANIVLDDADIDELIMPAYOMAMINSQOTCSALTRLIVPALITE -AE 298
OY 303 KISEVNAK-LSTGDFPD-NATVTPYVIDNKSADPIESLVDADAOKA-----KELNEFR 354
Db 299 TAAKTIAEAYTVGAPDDPDTTIGPLVSATQLRVRGYIDRGVQEGATLTITGSEPEVEGLA 358
OY 355 DGRLLTGLFDHVTLDPMKLAMEEPFGPIPIIRVKDAEBAVALANSKDPGLQSSVPTRDF 414
Db 359 VGYYKPIPIFSSVTYDMDTTHREIIGCPVSLAPDYDTEEAVALINDSEELGKGVWSRV 418
OY 415 QKAFDIANKLEVGTVHINNKTRGPDNPEFLGKSGAGVQGIKRSIEAMTVKSI 470
Db 419 DRARAARMRFGQWING--GEFNPNAPFGVKGSGTGRBGTGHLEFLFIKSL 472

RESULT 67
US-10-129-518-12
; Sequence 12, Application US/10129518
; Publication No. US20030215930A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Genes Involved in Cycloodecanone Degradation Pathway
; FILE REFERENCE: BC1023 PCT
; CURRENT APPLICATION NUMBER: US/10/129,518
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/170,214
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Rhodococcus ruber
US-10-129-518-12

Query Match      28.1%; Score 672; DB 15; Length 474;
Best Local Similarity 35.5%; Pred. No. 3,4e-45;
Matches 169; Conservative 84; Mismatches 197; Indels 26; Gaps 9

OY 8 YNNGEMKSVNO--TEIISPIDSSIGFVPAMTREBEVDHAKRAGREALPAMAALTYVERA 65
Db 10 YINGSWVASTSKTIVLEINPATBEVIGTVPDGTADVDAAVAARAAPFGWASTPDKXA 69
OY 66 QYLRK-AADIIERDKBEIATVTLAKETSKAYNASTVEVETADLIRYAEEGRILSTADE 124
Db 70 QYLRAAAGIADRSEDLARTTISAEMGAPLSFQAOMQVPLPINSFSHAAVAESFPFERT 129
OY 125 GKKMDASGKHLAVIRROPVGIVLAPRYNPNVLSGSLIARALIGNNVMPKPTQGSV 184
Db 130 GS-----SVTVEREPIGVGAITPMNYPPLHQIAAVALAAGNTIVKPSVADL 179
OY 185 SGLVLAKAPAEAGLDPAGVFNITITGSGSEIGDYIVHEEYVNFINTGSPVGORIGLAG- 243
Db 180 NAMMELIIDAAGVAPAGVFNVLVSGRPVVGELASHHEVDMLSTFQSTNAKGRVSLAAQ 239
OY 244 -MRPLMELGCKDAGIVTADADLDLNAKOIVAGAYDYSQQRCTAIRVILVEEVADELAE 302
Db 240 TVKRVALEIGKGSANIVLDDADIDELIMPAYOMAMINSQOTCSALTRLIVPALITE -AE 298
OY 303 KISEVNAK-LSTGDFPD-NATVTPYVIDNKSADPIESLVDADAOKA-----KELNEFR 354
Db 299 TAAKTIAEAYTVGAPDDPDTTIGPLVSATQLRVRGYIDRGVQEGATLTITGSEPEVEGLA 358
OY 355 DGRLLTGLFDHVTLDPMKLAMEEPFGPIPIIRVKDAEBAVALANSKDPGLQSSVPTRDF 414
Db 359 VGYYKPIPIFSSVTYDMDTTHREIIGCPVSLAPDYDTEEAVALINDSEELGKGVWSRV 418
OY 415 QKAFDIANKLEVGTVHINNKTRGPDNPEFLGKSGAGVQGIKRSIEAMTVKSI 470
Db 419 DRARAARMRFGQWING--GEFNPNAPFGVKGSGTGRBGTGHLEFLFIKSL 472

```

Oy		125	GCKMDASGHTGLATIRKQPVGIYALAIAPYNFPVNLSGSKIAIPALLIGVNWFKPPTQSSV	184
Db		130	GS-----SVIAREPIGVVALITPMNVLHQLAKVAAYLAAGNTIVVKSEVAPL	179
Oy		185	SGLVIAKAFAAGACPAVGFTNTITRGSEISGYIVEHEEVNFINTFGSTIPVQRITCKLAG-	243
Db		180	NAMMLAEIIIDAGCPACGFENIVSGTGPVGGALASHHEVDNISFTGSIINACKRVSLEAAQ	239
Oy		244	-MRPIMTELGKDAIGVIADADDLDNAAKQIYAGAIVDSGRCCTAIRVLVVEEVADEIAE	302
Db		240	TVEKEVALETGKGSANIVILDADIDIBELMPNAQVMAMINSQTCSALTLLVPRAILITE-AE	298
Oy		303	KISENVAK-LSVGPBPD-NATVTEVIDNSADPIESLIYVDAROKCA-----KEINPKR	354
Db		299	TAAKTIAAAYVGAAPDDPDTTLGLPVSAATOLKRVRGYIDRGVOBSATLLITGSEFEVGLA	358
Oy		355	DGRLLTPGLPFHVLTDMKLAMEEPFPGPLPIIRVWAEEAVALANKSDPGLQS SVETPDF	414
Db		359	VGYVVKPTIFSEVLPDMTIHREEIFGPLSLIAPYDTEBDARIANDSSEYLRGCVWSNDV	418
Oy		415	QKAFDIANKLEVGVVHHNNKTGRGEEDNEPFLGLKSGCAGVCGIKRYSIEAMTNVYSI	470
Db		419	DRARAIVAARMRTQGVMTING--GEFNPNAPFPGYGXQSGTGREFGTGHGLEFIKISL	472
 RESULT 68				
		US-10-156-761-11332		
		/ Sequence 11332, Application US/10156761		
		/ Publication No. US20030119018A1		
		/ GENERAL INFORMATION:		
		/ APPLICANT: OMCURA, SATOSHI		
		/ APPLICANT: IKEDA, HARUO		
		/ APPLICANT: ISHIKAWA, JUN		
		/ APPLICANT: HORIKAWA, HIROSHI		
		/ APPLICANT: SHIBA, TADAYOSHI		
		/ APPLICANT: SAKAKI, YOSHIOYUKI		
		/ APPLICANT: HATTORI, MASAHIRA		
		/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES		
		/ FILE REFERENCE: 249-262		
		/ CURRENT APPLICATION NUMBER: US/10/156, 761		
		/ CURRENT FILING DATE: 2002-05-29		
		/ PRIOR APPLICATION NUMBER: JP 2001-204089		
		/ PRIOR FILING DATE: 2001-05-30		
		/ PRIOR APPLICATION NUMBER: JP 2001-272697		
		/ PRIOR FILING DATE: 2001-08-02		
		/ NUMBER OF SEQ ID NOS: 15109		
		/ SEQ ID NO 11332		
		/ LENGTH: 499		
		/ TYPE: PRF		
		/ ORGANISM: Streptomyces avermitilis		
		US-10-156-761-11332		
		Query Match	28.1%; Score 671.5; DB 14; Length 499;	
		Beet Local Similarity	36.2%; Pred. No. 4e-45;	
		Matches 175; Conservative 84; Mismatches 194; Indels 31; Gaps 12;		
Oy		8	YVNGEWKS-SVNOJFILSPIDDSSIGFVPANTREBYDHAMAAGRALPAMMAALTYBERAQ	66
Db		24	FIGESWPAFDGHHYEVIDPATEAVVGMPEASRDQVHAANAARAFAAMSSTPAEBEAA	83
Oy		67	YLHKAADIIEEDKEEIATVLAKEISKAVNASVTEVVR-----TADLIRYA---AEEG	115
Db		84	VLGRAAEWRNRNLYPYA-----ELAQAESGATTGARAKMOVGAGAAPRRRIARVPAEPP	138
Oy		116	IRLSTADEGSKMD-ASTGHLAVIIRQPVGIYALAIAPYNFPVNLSGSKIAIPALLIGNVV	174
Db		139	I--AQQINEABPFGGAAYMGALAV--RPVGVVTICTSYNNPWNPAPACKIAPALAMGNTV	194
Oy		175	MFKPPTQSSVGLVLAKAFAAGCPACGFNTITRGSSITGYIYEHDEVNFINTFGSTIPV	234
Db		195	VVKRAPODPLSVYRNAEALQAAGVPPGVNVVSGSQAQAGAAADVCDPVDWVSFTGSAIV	254


```
RESULT 71
US-10-369-493-7213
; Sequence 7213, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7213
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-369-493-7213

Query Match          27.9%; Score 668; DB 15; Length 468;
Best Local Similarity 34.2%; Pred. No. 7e-45;
Matches 162; Conservative 82; Mismatches 199; Indels 30; Gaps 10;

QY 10 NQEMSSVNVQIILIPIDSSIGFVPAMTREVVDHAKMGREAL--PAMALTYTERAOY 67
DB 3 NQESSEBIN-----PIDSSVAATLAQASESDVDAVRAARQALETPARDMKVHERGRL 56
QY 68 LHKADIIERDKERIATVLAKEISKAYNASVTEVTRADLIRYAAEBGIRLSTDSDEGK 127
DB 57 LHRFDLIDQNFETLAQITRNGKITISRRQSKASADMFYFA-----AVCETTES 110
QY 128 MDASTGHKLAVIRROPVGIIVLAIPYVNVNLSGSKIAPALIGNVNMEKPTQGSVSG 187
DB 111 VPSRGVNFCSFSPFVGVVAIAIPWNSPITLEAQGLAPALAAAGNAIVLKSEVTPCIGL 170
QY 168 VLAKEAFARGLPAGVFNITITGRSGEIGDYIVHEEVNFNFTGSGTPVQGRIGKLAMR-- 245
DB 171 QYGRNAMEGFPPIGLNVVITGFGASVGRALVAHPGDVMTFGCTKSGREIAKVAGERLL 230
QY 246 PIMLEIGGKADGIYVADADLDNAKOIVAGAYDVSQORCTAIKRYLVVEEVDLAEKIS 305
DB 231 PCILEIGGKSPNIVFDADIDQAVTGMLEFISNMGQSCIASRILIDQGIYDAVARKLT 290
QY 306 ENVAKLISVGDPPD-NATVTPVIDNSADPIESLIVDAROKAKEINEFPR-DGRLLTPGL 363
DB 291 SATRALKVSGPYDADJVAAPVSGFPHRDIHEHVERALDSQAVALCGSGRPSGGILDKGA 350
QY 364 FDHVTL-----DMKLAMEEPFGPILPIIRYKDAEBAVAIANKSDFGLOSVFTDFOKA 417
DB 351 YVQPIILEVTDERSVAHEIFGPVACWKFQDEBDLYRIANNYALGACGWTADYKKA 410
QY 418 FDIANKLEVGVTHINNKGRGPDNFPPIGLKSGS-----AGVQGI-RYSIEANT 465
DB 411 FRAARIRAKGVWNTYK-IAEVNVPFGGFGKSGLAGREGCIQGMREIMLEKST 462

RESULT 72
US-09-815-242-5058
; Sequence 5058, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
```

```
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5058
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5058

Query Match          27.9%; Score 668; DB 9; Length 483;
Best Local Similarity 33.2%; Pred. No. 7.3e-45;
Matches 158; Conservative 90; Mismatches 212; Indels 16; Gaps 7;
```

```
QY 6 QNVYNGEKSQVNV--QIEILSPIDSSIGFVPAMTREVVDHAKMGREALPAMALTYVE 63
DB 12 QAVYDGAWVDNDNGCTIVNNPATGETIGSVPKGMAETRPATLADALPAMRLTAKE 71
QY 64 RAOYLHKADIIERDKERIATVLAKEISKAYNASVTEVTRADLIRYAAEBGIRLSTDSAD 123
DB 72 RANKLRMFPDLMINQDRLARLMTIEQKPLAEAKGEIAYVASFLEWGEAKIYGTI 131
QY 124 EGGKMDASTGHKLAVIRROPVGIIVLAIPYVNVNLSGSKIAPALIGNVNMEKPTQGS 183
DB 132 PCHQPD-----KRIIVIKQPIGVTAALTPMNPSPAMITRKAGPLAAGCTVTLKPSQTP 186
QY 184 VSGVLAKAFABAGIPACVFNITITGRSGEIGDYIVHEEVNFNFTGSGTPVQGRIGK--L 241
DB 187 YSALALAEIARAGIPKGVFSVVTGSAGEVGGELTSPYVKLTFTGSTETIGRQUMECA 246
QY 242 AGMPIMLEIGGKADGIYVADADLDNAKOIVAGAYDVSQORCTAIKRYLVVEEVDLAEL 301
DB 247 QDIKVSIELEGNNAPFIVFDADIDDAVEGALISKYRNNGTCCANRLTYQDDGYDAFV 306
QY 302 EKISENVAKISVGDPPD-NATVTPVIDNSADPIESLIVDAROKAKEINEF--RDGR 357
DB 307 DKLKAAYAKLNIINGLEIGVTTGPLIDAKAVAKYEHIADAVSKAGVSGSKPHALGCT 366
QY 368 LITPGLPHTVTLDMKLAMEEPFGPILPIIRYKDAEBAVAIANKSDFGLOSVFTDFOKA 417
DB 367 FPEFTIIVDVKNALVSKDETFGLAPVFRKDAEVAIAMSNDTEFGIASYFVADLARKV 426
QY 418 FDIANKLEVGVTHINNKGR-GPDNFPPIGLKSGGAGVQGI-RYSIEANTVKSIVL 472
DB 427 FRVAQGLEYGAVGIN--TGLISNEVAPFGGIKAGSGLREGSKYGLIEDVLEIKYICL 480

RESULT 73
US-10-282-122A-43455
; Sequence 43455, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
```


Db 250 ELGGAPFLVLDADDDDAVEAAAFGAYFNQOICMSTER-LIVTAVAAPFVEKLAARKVA 308
310 KLSVGDPPFNATVY--PVIDNNSADFIESTLVDAROKAGELMBFKRDGRLTGGFDHYT 368
Db 309 TIRAGDPNDPOSVLGLIDANAGQRIQVLDLALAGAAQVVGCGIGIMPPMLDQYT 368
369 LDMKLAMEPFGPIPIIRKDAEBAVALANKSDFGLOSSVFTRDFOKAFDIANKLEVGT 428
369 EEMRLYREESFGPAVAVLKGDDDEBLRLANDSEFGLSAIFSRDVSRAAMELAQRVDSGI 428
QY 429 VHINKTGRGPDNFPPLGLKSGAGVQIGIRYSIEANTNVKSIYL 472
Db 429 CHINGPTVHDEAQMFGGVKSSGYSGFSGRASIEHFTQLRWLTI 472

RESULT 75

US-09-750-986D-28
; Sequence 28, Application US/09750986D
; Publication No. US20030228670A1
; GENERAL INFORMATION:
; APPLICANT: Steinhuchel, Alexander
; APPLICANT: Priefert, Horst
; APPLICANT: Rabenhorst, Jürgen
; TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF CONIFERYL
; TITLE OF INVENTION: ALCOHOL, CONIFERYLALDEHYDE, FERULIC ACID, VANILLIN AND
; TITLE OF INVENTION: VANILLIC ACID AND THEIR USE
; FILE REFERENCE: Bayer-9998-CAO
; CURRENT APPLICATION NUMBER: US/09/750,986D
; CURRENT FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 196 49 655.1 GERMANY
; PRIOR FILING DATE: 1996-11-29
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Pseudomonas sp.
US-09-750-986D-28

Query Match 27.9%; Score 667.5; DB 10; Length 481;

Best Local Similarity 36.4%; Pred. No. 8e-45;

Matches 169; Conservative 80; Mismatches 202; Indels 13; Gaps 6;

QY 15 SSVNOIEISPIDDSLGFVPMATREEVHAMKAGREALPAWALTVYERAOYLHKADI 74
Db 16 SDEFTFERSSPLTGEVSVVAAASLEDDAVALAAQAAPFEMALAPSERARILRAADI 75
QY 75 IERDKREIATVLAKELISKAYNASVTEVVTADLIRYAAEGIRLSTADEGGRMDASTGH 134
Db 76 IER-DRSSFTAAASFTGAAGNMYGFNVYLAAGMLREA-----AMTTQIQGVIPENVP 129
QY 135 KLAVIRROQVGIVLAAPYNYPVNLSGKIAPALLIGANNVMPPTQGSVSGVLAKAPA 194
Db 130 SFAMAVRQPCGVVLGIAPNNAPVILGVRVAVPLACGNTVVLKSSSELSPFTHRLIGQVLA 189
QY 195 BAGLPAGVNTITGRSE---IGDYVBEHEVNFITGSTPYGQRIKLAG--MRPIML 249
Db 190 DAGLGDGVNVVSNAPQDPAVERLIANPAVRVNFSTGHVGRITIGELISARHLKPAVL 249
QY 250 ELGGKADAGIVLADADLDNAKQIVAGAYDSGQRCYAIRVLAVEREVADELAEKISENVA 309
Db 250 ELGGKAPFLVLDADDDDAVEAAAFGAYFNQOICMSTER-LIVTAVAAPFVEKLAARKVA 308
QY 310 KLSVGDPPFNATVY--PVIDNNSADFIESTLVDAROKAGELMBFKRDGRLTGGFDHYT 368
Db 309 TIRAGDPNDPOSVLGLIDANAGQRIQVLDLALAGAAQVVGCGIGIMPPMLDQYT 368
QY 369 LDMKLAMEPFGPIPIIRKDAEBAVALANKSDFGLOSSVFTRDFOKAFDIANKLEVGT 428
Db 369 EEMRLYREESFGPAVAVLKGDDDEBLRLANDSEFGLSAIFSRDVSRAAMELAQRVDSGI 428
QY 429 VHINKTGRGPDNFPPLGLKSGAGVQIGIRYSIEANTNVKSIYL 472

Db 429 CHINGPTVHDEAQMFGGVKSSGYSGFSGRASIEHFTQLRWLTI 472

Search completed: August 28, 2005, 10:40:14
Job time : 169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2005, 10:22:14 ; Search time 43 Seconds
(without alignments)
824.612 Million cell updates/sec

Title: US-09-868-195-12

Perfect score: 2393

Sequence: 1 LTKKQYVNGWKKSSVNOI.....GIRSIEMTNKSTVLDNK 475

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 80 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/prodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Query Length	DB ID	Description
1	1808.5	75.6	474	US-09-583-110-3808	Sequence 3808, Ap
2	1692	70.7	436	US-09-107-433-3804	Sequence 3804, Ap
3	786.5	32.9	492	US-09-717-926-4	Sequence 4, Appl
4	775.5	32.4	492	US-09-634-955B-13	Sequence 13, Appl
5	733.5	30.7	526	US-09-552-991A-27302	Sequence 27302, A
6	711	29.7	487	US-09-489-039A-12633	Sequence 12633, A
7	705.5	29.5	491	US-09-585-174-3	Sequence 1, Appl
8	672	28.1	474	US-09-732-615-12	Sequence 12, Appl
9	672	28.1	474	US-10-273-051-12	Sequence 12, Appl
10	668	27.9	594	US-09-252-991A-27749	Sequence 27749, A
11	667.5	27.9	481	US-08-976-063E-28	Sequence 28, Appl
12	660	27.6	517	US-09-538-092-1047	Sequence 1047, Ap
13	659.5	27.6	508	US-08-956-171E-5241	Sequence 5241, Ap
14	659.5	27.6	508	US-08-781-986A-5241	Sequence 5241, Ap
15	658	27.5	486	US-09-328-352-4767	Sequence 4767, Ap
16	658	27.5	500	US-09-538-092-831	Sequence 831, App
17	650	27.2	497	US-09-602-777A-182	Sequence 182, App
18	650	27.2	512	US-09-538-092-1119	Sequence 1139, Ap
19	650	27.2	512	US-09-961-403-12	Sequence 12, Appl
20	645	27.0	521	US-09-321-294-2	Sequence 2, Appl
21	642.5	26.8	507	US-09-489-039A-12722	Sequence 12722, A
22	639	26.7	532	US-09-949-016-9207	Sequence 9207, Ap
23	635	26.5	498	US-09-489-039A-7569	Sequence 7569, Ap
24	632.5	26.4	505	US-09-543-681A-7151	Sequence 7151, Ap
25	632	26.4	496	US-09-543-681A-5701	Sequence 5701, Ap
26	631	26.4	529	US-09-252-991A-23785	Sequence 23785, A
27	626.5	26.2	505	US-09-328-352-5823	Sequence 5823, Ap

28	626	26.2	502	4	US-09-540-236-3780	Sequence 3780, Ap
29	625.5	26.1	510	4	US-09-252-991A-17138	Sequence 17138, A
30	624	26.1	510	3	US-09-134-001C-4541	Sequence 4541, Ap
31	623	26.0	493	4	US-09-328-352-6753	Sequence 6753, Ap
32	621	26.0	485	3	US-09-651-941-9	Sequence 9, Appl
33	621	26.0	485	4	US-09-955-597-9	Sequence 9, Appl
34	621	26.0	508	3	US-09-655-270A-9	Sequence 9, Appl
35	621	26.0	517	4	US-09-538-092-851	Sequence 851, App
36	621	26.0	520	4	US-09-949-016-10653	Sequence 10653, A
37	618	25.8	502	4	US-09-328-352-6687	Sequence 6687, Ap
38	616.5	25.8	439	4	US-09-949-016-9559	Sequence 9559, Ap
39	616	25.7	482	3	US-09-155-183-4	Sequence 4, Appl
40	616	25.7	482	4	US-09-733-383-4	Sequence 4, Appl
41	611.5	25.6	514	4	US-09-710-279-2682	Sequence 2682, Ap
42	609	25.4	497	1	US-08-513-841-2	Sequence 2, Appl
43	609	25.4	497	2	US-08-696-834-2	Sequence 2, Appl
44	609	25.4	497	2	US-08-942-673-2	Sequence 2, Appl
45	609	25.4	497	3	US-09-118-317-2	Sequence 2, Appl
46	607.5	25.4	518	3	US-09-134-001C-4451	Sequence 4451, Ap
47	607	25.4	498	4	US-09-344-882-22	Sequence 22, Appl
48	604.5	25.3	498	4	US-09-328-352-5847	Sequence 5847, Ap
49	602	25.2	518	4	US-09-976-594-287	Sequence 287, App
50	602	25.2	518	4	US-09-919-039-143	Sequence 143, App
51	602	25.2	553	4	US-09-949-016-7505	Sequence 7505, Ap
52	601	25.1	496	4	US-09-710-279-1030	Sequence 1030, Ap
53	601	25.1	506	3	US-09-134-001C-4383	Sequence 4383, Ap
54	600.5	25.1	514	4	US-09-489-039A-13820	Sequence 13820, A
55	599	25.0	493	4	US-09-538-092-1147	Sequence 1147, Ap
56	598	25.0	490	4	US-09-949-016-5979	Sequence 5979, Ap
57	598	25.0	490	4	US-09-328-352-6118	Sequence 6118, Ap
58	597.5	25.0	423	4	US-09-717-926-2	Sequence 2, Appl
59	595.5	24.9	477	4	US-09-902-540-15889	Sequence 15889, A
60	594	24.8	505	4	US-09-328-352-7337	Sequence 7337, Ap
61	591.5	24.7	748	4	US-09-252-991A-28449	Sequence 28449, A
62	590.5	24.7	500	4	US-09-328-352-6757	Sequence 6757, Ap
63	585	24.4	499	4	US-09-543-681A-4515	Sequence 4515, Ap
64	583.5	24.4	488	3	US-09-134-001C-4245	Sequence 4245, Ap
65	581.5	24.3	492	4	US-09-489-039A-9023	Sequence 9023, Ap
66	578	24.2	502	4	US-09-489-039A-13971	Sequence 13971, A
67	577.5	24.1	481	4	US-09-328-352-4475	Sequence 4475, Ap
68	568.5	23.8	486	4	US-09-344-882-20	Sequence 20, Appl
69	562.5	23.5	458	4	US-09-716-865-6	Sequence 6, Appl
70	562	23.5	475	4	US-09-252-991A-32298	Sequence 32298, A
71	561	23.4	642	4	US-09-252-991A-25398	Sequence 25398, A
72	559.5	23.4	534	4	US-09-344-882-24	Sequence 24, Appl
73	555	23.2	535	4	US-09-538-092-1173	Sequence 1173, Ap
74	555	23.2	564	4	US-09-328-352-7905	Sequence 7905, Ap
75	554.5	23.2	468	4	US-09-351-150A-9	Sequence 9, Appl
76	553.5	23.1	493	3	US-09-134-001C-4388	Sequence 4388, Ap
77	552	23.1	477	4	US-09-328-352-6549	Sequence 6549, Ap
78	551.5	23.0	479	4	US-09-710-279-2956	Sequence 2956, Ap
79	550.5	23.0	692	4	US-09-252-991A-26724	Sequence 26724, A
80	539.5	22.5	498	4	US-09-648-004-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-583-110-3808
; Sequence 3808, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATHO0-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12

PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO: 3808
LENGTH: 474
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-3808

Query Match 75.6%; Score 1808.5; DB 4; Length 474;
Best Local Similarity 73.7%; Pred. No. 4,7e-161;
Matches 350; Conservative 59; Mismatches 65; Indels 1; Gaps 1;

QY 1 LTRKEQYNYNGEMKSSVNOIETLSPIDSSLGFPVPMREEDVHMKAREALPMAALT 60
DB 1 LTR-YQNLVNGKMKSSQEGITITSPINOELGTVPMTOTEDAMQARALPMAKALS 59
QY 61 YVERAQYLHKAADIIRDKKEIATYLAKEISKAYNASVTEVVRTADLIRYAEGRIST 120
DB 60 AIERAAVYHKTALIERDKKEIGTILAKEVAKIKAKIAGEVVRTADLIRYAEGRITG 119
QY 121 SADEGKMDASTGHKLAIVRRQPGIVLAIAPVNPNVLSGSKIPALIGVNMFKPPT 180
DB 120 QMEGGGFEAASKNKLAIVRRPVGIVLAIAFPNPNVLSASKIAPALIAGVNMFKEPT 179
QY 181 QGSVGLVLAKAFAPAGLIPAGVNTITGRSEIGDIYVEHEVNFINTGSTPGVQRIK 240
DB 180 QGSISGLLAKAFERAGIPAGVNTITGRSEIGDIYIHKVNFINTGSTPIGERICR 239
QY 241 IAGMRPIMELGKDAIVLADLDNNAKOIVAGAYDYSGORCTAIKRVLVEEVADEL 300
DB 240 IAGMRPIMELGKDAIVLEDDLEHAKOIVAGAFSYSGORCTAIKRVILVESVADKL 299
QY 301 AEKISENVAKLSVGPFDNATVTPVIDNSADPIESLVYDAROKAKELNEKRGRLIT 360
DB 300 ATLLEBVSCLTVGDPFNADITPVINDASADPIWGLIEDAQEKQALTPIKREGLTW 359
QY 361 PGLFPHVTLDMKLAEEFPGLPIPIIRYKDAEAVAIANKSDFGLQSSVFTRDFOKAPDI 420
DB 360 PGLFQVTKDMKVAMBEFPGLPIPIIRVASVEEALAFANESRFGIQQSSVFTDFOKAPDI 419
QY 421 ANKLEVGTVHINNTKRGPDNFPFLGLKSGAGVQIGIRYSIEMTNVKSIVLDMK 475
DB 420 AEKLEVGTVHINNTKRGPDNFPFLGVKSGAGVQIGIRYSIEMTNVKSIVPDVK 474

RESULT 2

US-09-107-433-3804
Sequence 3804, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID and AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131

FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Atinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3804:
SEQUENCE CHARACTERISTICS:
LENGTH: 436 amino acids
TOPOLOGY: linear
TYPE: amino acid
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...436
SEQUENCE DESCRIPTION: SEQ ID NO: 3804:
US-09-107-433-3804

Query Match 70.7%; Score 1692; DB 4; Length 436;
Best Local Similarity 74.9%; Pred. No. 3.6e-150;
Matches 326; Conservative 53; Mismatches 56; Indels 0; Gaps 0;

QY 41 EVDHAKAKAREALPMAALTVEERAOYLHKAADIIRDKKEIATYLAKEISKAYNASYTE 100
DB 2 EADEMQARALPMAKLSAIBERAAVYHKTALIERDKKEIGTILAKEIKAKIAGE 61
QY 101 VVRTADLIRYAEGRISTSADEGKMDASTGHKLAIVRRQPGIVLAIAPVNPNVLS 160
DB 62 VVRTADLIRYAEGRITRTQAMEGGFEAASKNKLAIVRRPVGIVLAIAFPNPNVLS 121
QY 161 GSKIPALIGVNMFKPPTQGSVGLVLAKAFAPAGLIPAGVNTITGRSEIGDIYEH 220
DB 122 ASKIPALIAGVNMFKPPTQGSISGLLAKAFERAGIPAGVNTITGRSEIGDIYIEH 181
QY 221 EEVNFINTGSTPGVQRIKLAGMRPIMELGKDAIVLADLDNNAKOIVAGAYDYS 280
DB 182 KEVNFINTGSTPIGERIGRLAGMRPIMELGKDAIVLEDDLEHAKOIVAGAFSYS 241
QY 281 GORCTAIRVLEVEVADELAEKISENVAKLSVGPFDNATVTPVIDNSADPIESLVVD 340
DB 242 GORCTAIRVILVESVADKLATLQEBVSCLTVGDPFNADITPVINDASADPIWGLIED 301
QY 341 AROKAKELNFPKRGRLITPGLFDHVTLDMLAMEEFPGLPIPIIRYKDAEAVAIANK 400
DB 302 AOEKAOALTPIKREGLNLTWVLPDQYTKDMKVAMBEFPGLPIPIIRVASVEEALAFANE 361
QY 401 SDFGLQSSVFTRDFOKAPDIANKLEVGTVHINNTKRGPDNFPFLGLKSGAGVQIGIRYS 460
DB 362 SEFGLQSSVFTNDFOKAPDIANKLEVGTVHINNTKRGPDNFPFLGVKSGAGVQIGIRYS 421
QY 461 IEAMTNVKSIVLDMK 475
DB 422 IEAMTNVKSIVPDVK 436

RESULT 3

US-09-717-926-4
Sequence 4, Application US/09717926
Patent No. 6569657
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel
APPLICANT: Cook, William J.
TITLE OF INVENTION: 32140, A No. 6569657e1 Human Aldehyde
TITLE OF INVENTION: Dehydrogenase
FILE REFERENCE: 35800/205243
CURRENT APPLICATION NUMBER: US/09/717, 926

```
/ CURRENT FILING DATE: 2000-11-21
/ PRIOR APPLICATION NUMBER: US 60/214,707
/ PRIOR FILING DATE: 2000-06-27
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 492
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Aldehyde dehydrogenase Pfam consensus sequence
US-09-717-926-4

Query Match      32.9%; Score 786.5; DB 4; Length 492;
Best Local Similarity 38.6%; Pred. No. 4,5e-65;
Matches 193; Conservative 92; Mismatches 166; Indels 49; Gaps 16;

QY 12 EWKSVN--QIEILSPIDSS--LGFVPAMTREVDHAKAGREAL--PAMALTYVERA 65
DB 1 EWDVDSAGKTFEVNPAKNGEVIIGRVPEATIEDVDAAVAAKAEAFKSGPMKVPASERA 60
QY 66 QYIHKAAIDIERDEKEIATVLAKEISK--AYNASVTEVVRTADLIRYAA-----EEG 115
DB 61 RILRLKADLIERDEDELALETLDGKPLAEAKGTEVGRALDEIRYAGMARKLMGERR 120
QY 116 I--RLSTADGCKMDASTGHLAVIRROPVGIYLAIPYNPVNLSSGIAPALIGNV 173
DB 121 VLPSLATDGE-----ELNYYTRREPGLGVGVISPMNPLALMLKLAIPALAGNT 170
QY 174 VMEKPTQGSVSGVLAKAPAEAG--LPAGVFNTITGRSGEIGDYVEHEEVNFINFTG 230
DB 171 VVLRKSEQPTPLALLAEILIEEGANNLPKGVNVVPGGAEVGOALLSHPIDKISFTG 230
QY 231 STPVGQRIKLA---GMRPIMLELGKDAIGIYLAADLDNAKQIYAGAYDYSGRCTAI 287
DB 231 STEVGLIMEAANAANKLKVITIELGKSPVIYFDADLDKAVRIVFGAFGAGVCICAP 290
QY 288 KAVLVEEVADELAKEISENVAKLS--VGDPFNAT--VTPVIDNSADPIESLAVDAROK 344
DB 291 SKLHVHESIYDEVEKLERVKKLKLIIDPLSDTNIYGPLISEQGFDPVLSTIEDGKEE 350
QY 345 GAK-----ELNEFKRDGRLITPGLFDHVTLDMLAMEBPPGRLPIIRYKDAEAVA 396
DB 351 GAKVLCGERDSKEYLGGGYVQPTIFDVPDKIMKEELFGVPLPIIKFDLDEALE 410
QY 397 IANKSDPGLQSSVFTPD--FOKAFDIANKLEVGTVINNTGGRPD--NFPPLGLK--GSGAG 453
DB 411 LANDREYGLAAVFTFDILARARVAKALEAGIWNVNDVCVAHAEPQLRFFGVKQSSGIG 470
QY 454 VQ--GIRYSIEAMTNVKSIVL 472
DB 471 REHGKYGLEEYETIKVTI 490

RESULT 4
US-09-634-955B-13
/ Sequence 13, Application US/09634955B
/ Patent No. 6511834
/ GENERAL INFORMATION:
/ APPLICANT: Meyers, Rachel
/ TITLE OF INVENTION: 32142, 21481, 25964, 21686, NOVEL HUMAN DEHYDROGENASE
/ FILE REFERENCE: MNI-134
/ CURRENT APPLICATION NUMBER: US/09/634,955B
/ CURRENT FILING DATE: 2000-08-08
/ PRIOR APPLICATION NUMBER: 60/192,002
/ PRIOR FILING DATE: 2000-03-24
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: FastSeq for Windows Version 4.
/ SEQ ID NO 13
/ LENGTH: 492
/ TYPE: PRT
```

```
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Aldehyde dehydrogenase family domain
US-09-634-955B-13

Query Match      32.4%; Score 775.5; DB 4; Length 492;
Best Local Similarity 38.4%; Pred. No. 4,9e-64;
Matches 192; Conservative 92; Mismatches 167; Indels 49; Gaps 16;

QY 12 EWKSVN--QIEILSPIDSS--LGFVPAMTREVDHAKAGREAL--PAMALTYVERA 65
DB 1 EWDVDSAGKTFEVNPAKNGEVIIGRVPEATIEDVDAAVAAKAEAFKSGPMKVPASERA 60
QY 66 QYIHKAAIDIERDEKEIATVLAKEISK--AYNASVTEVVRTADLIRYAA-----EEG 115
DB 61 RILRLKADLIERDEDELALETLDGKPLAEAKGTEVGRALDEIRYAGMARKLMGERR 120
QY 116 I--RLSTADGCKMDASTGHLAVIRROPVGIYLAIPYNPVNLSSGIAPALIGNV 173
DB 121 VLPSLATDGE-----ELNYYTRREPGLGVGVISPMNPLALMLKLAIPALAGNT 170
QY 174 VMEKPTQGSVSGVLAKAPAEAG--LPAGVFNTITGRSGEIGDYVEHEEVNFINFTG 230
DB 171 VVLRKSEQPTPLALLAEILIEEGANNLPKGVNVVPGGAEVGOALLSHPIDKISFTG 230
QY 231 STPVGQRIKLA---GMRPIMLELGKDAIGIYLAADLDNAKQIYAGAYDYSGRCTAI 287
DB 231 STEVGLIMEAANAANKLKVITIELGKSPVIYFDADLDKAVRIVFGAFGAGVCICAP 290
QY 288 KAVLVEEVADELAKEISENVAKLS--VGDPFNAT--VTPVIDNSADPIESLAVDAROK 344
DB 291 SKLHVHESIYDEVEKLERVKKLKLIIDPLSDTNIYGPLISEQGFDPVLSTIEDGKEE 350
QY 345 GAK-----ELNEFKRDGRLITPGLFDHVTLDMLAMEBPPGRLPIIRYKDAEAVA 396
DB 351 GAKVLCGERDSKEYLGGGYVQPTIFDVPDKIMKEELFGVPLPIIKFDLDEALE 410
QY 397 IANKSDPGLQSSVFTPD--FOKAFDIANKLEVGTVINNTGGRPD--NFPPLGL--KQSGAG 453
DB 411 LANDREYGLAAVFTFDILARARVAKALEAGIWNVNDVCVAHAEPQLRFFGVHQSIGIG 470
QY 454 VQ--GIRYSIEAMTNVKSIVL 472
DB 471 REHGKYGLEEYETIKVTI 490

RESULT 5
US-09-252-991A-27302
/ Sequence 27302, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 27302
/ LENGTH: 526
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27302

Query Match      30.7%; Score 733.5; DB 4; Length 526;
Best Local Similarity 36.8%; Pred. No. 4,8e-60;
Matches 172; Conservative 87; Mismatches 186; Indels 23; Gaps 7;

QY 11 GEMK--SSVNOIEILSPIDSSLGFVPAMTREVDHAKAGREALPAMALTYVERA 68
```

Db 53 GEMRGRGRRLKVNPNPDGSLLEBEQADRDLDLAAVAKAEVOPAWAALPPSAAAVL 112
QY 69 HKAADIIEKDEKEIITVLAKEISKAYNASVTEVVRADLIRYAEEGISTLSA----DE 124
Db 113 YKAVEFDRRHEIYDWIIRRESG-----TRUKAEIENGARAILTESASFPARVH 163
QY 125 GGRKMDASTGHLKAVIRROPVGIVLAIPYNYPVNLSGSKIAPALIGNVVMEKPPPTGVS 184
Db 164 GRIVSDVPGKESRYRSALIGVGVISFPWNPFLLHQRIQRIAPALLAGNAVVKPAPSDP 223
QY 165 -SGVLAKAPAEAGLPAGVFNTITGRSGEIGDIYVHEHEVNINFTGSPVQORIGKLAG 243
Db 224 CGGLLARIIFEBAAGPAGLFSVYVGGSEIGDAFVEHPVPGLVTFGTSPVGRNIGRIAS 283
QY 244 ----NRPIMLEGKDGAGIVLADADLDNNAKQIVAGAYVSGORCAIKRVLVVEEVADE 299
Db 284 GGAHLKHYALELGSNPFVVLGDADLEQVNAAPVGEKFLHQICIMALNCITIEBLSYDA 343
QY 300 LAEKISENVAKLSVGP--FDNATVTPVIDNSADPIESLVVDAROKAKEINEFKRDR 357
Db 344 FAARFVERVKGLRVGDPQADTA-VGPIVNAQBLEGLKIKILARQEGAKPLIEGSGDQ 402
QY 358 LITPGLFDHVTLDMLAMEBPPGILPIIRVYDAEBAVAIAKNSDFGLSSVFTTRDPQKA 417
Db 403 ILAPVHFGVITATMEIARDEIFGPIVGLRLARDEAHLELANASEGLSSAVFSRDLERA 462
QY 418 FDIANKLEGVTHINNKTRGPDNPFPLGLKSGAGVQIGRISEAMT 465
Db 463 VAPAQQLRGMTHVNDIPVNDANAPFGEKNSGLGRFGMDWAIEBFT 510

RESULT 6
US-09-489-039A-12633
Sequence 12633, Application US/09489039A

Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709 2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 12633
LENGTH: 487
TYPE: PRN
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12633

Query Match 29.7%; Score 711; DB 4; Length 487;
Best Local Similarity 34.1%; Pred. No. 5, Se-58;
Matches 163; Conservative 90; Mismatches 209; Indels 16; Gaps 7;

QY 6 QNYVNGEMKSVN--QIETLSPIIDSLGFVPAMTEEEVDHAMKAGREALPAAALTYVE 63
Db 17 QAMIDGQMRDAPNGDVIAVTNPANGEOGSLVRKMGDETREAIERANRALPAMRALITAE 76
QY 64 RAQYIHKAAADIIEKDEKEIATVLAKEISKAYNASVTEVVRADLIRYAEEGISTLSAD 123
Db 77 RANILIRRMFDLMMENDDLARIIMTLEQKPLAEAKEISYASFIEMFAEGRKRIYDGTI 136
QY 124 EGGKMDASTGHLKAVIRROPVGIVLAIPYNYPVNLSGSKIAPALIGNVVMEKPPPTGS 183
Db 137 PEHQAD-----KRLVHKQPIGVTAITWNPAAAITRKAGPALAAGCTWALKPASQTP 191
QY 184 VSGVLAKAPAEAGLPAGVFNTITGRSGEIGDIYVHEHEVNINFTGSPVQORIGKLAG 243
Db 192 FSAALALAEIANRAGIPAGVFNVTGSAGAVGSELTSPVLRKISFGSIEIGRLMEQCA 251
QY 244 --MRPIMLEGKDGAGIVLADADLDNNAKQIVAGAYVSGORCTAIKRVLVVEEVADELA 301

Db 252 KDIKVSLELGNAPFIIFDDADLDKAVEGALASKFRNAGATCCANLTYQDGVYDRFA 311
QY 302 EKISENVAKLSVGP--DNATVTPVIDNSADPIESLVVDAROKAKEINEFK--RDR 357
Db 312 EKDQAAVEKLKIGDLDQDGVTTGFLIDEKAAKYBEHIALIANKAKVYTGKPHALG 371
QY 358 LITPGLFDHVTLDMLAMEBPPGILPIIRVYDAEBAVAIAKNSDFGLSSVFTTRDPQKA 417
Db 372 FPGQPIIVNVDSAKVAKEETFGPLAIFRKEADADVIAQANDPFGIAAFVARDLSRV 431
QY 418 FDIANKLEGVTHINNKTRGPDNPFPLGLKSGAGVQIGRISEAMTNYKSIYLD 474
Db 432 FRVGRALRYGIIGIN--TGIISTEVAPEGVAKSGLGREGSKYGIEDIETIKMCI 487

RESULT 7
US-09-585-174-3
Sequence 3, Application US/09585174

Patent No. 6586229
GENERAL INFORMATION:
APPLICANT: Ben-Bassat, Arie
APPLICANT: Catermole, Monica
APPLICANT: Gatenby, Anthony A.
APPLICANT: Gibson, Katherine J.
APPLICANT: Ramos-Gonzalez, Isabel
APPLICANT: Ramos, Juan
TITLE OF INVENTION: Method for the Production of p-Hydroxybenzoate in Species of
FILE REFERENCE: BCI018 US NA
CURRENT APPLICATION NUMBER: US/09/585,174
CURRENT FILING DATE: 2000-06-01
NUMBER OF SEQ ID NOS: 112
SOFTWARE: Microsoft Office 97
SEQ ID NO 3
LENGTH: 491
TYPE: PRN
ORGANISM: Pseudomonas mendocina KR-1
US-09-585-174-3

Query Match 29.5%; Score 705.5; DB 4; Length 491;
Best Local Similarity 34.5%; Pred. No. 1, Se-57;
Matches 167; Conservative 94; Mismatches 198; Indels 25; Gaps 8;

QY 5 YQNY-----VNGEWMK--SSVNOIETLSPIIDSLGFVPAMTEEEVDHAMKAGREALPAA 58
Db 8 YQNELQPLAGQWRAGSSGRPLEVFPDYNDELRLIALASREDDLAAVRKADSOREWAT 67
QY 59 LTVERRAQYLKAAADIIEKDEKEIATVLAKEISKAYNASVTEVVRADLIRYAEEGIRL 118
Db 68 TAPARARAVLEAVKIPERRREIYDWIIRRESG-----TRUKAEIENGARAILTESASFPARVH 118
QY 119 STSADG--GKMDAST-GHKLAVIRROPVGIVLAIPYNYPVNLSGSKIAPALIGNAV 174
Db 119 ESASIPNRVHGRILSNISNGKESRYRAPLGVIGISFPWNPFLLHQRIQRIAPALLAGNAV 178
QY 175 MEKPPPTGVS-GVLAAPAEAGLPAGVFNTITGRSGEIGDIYVHEHEVNINFTGSP 233
Db 179 VVKPASTPPIITGGLLARIIFEBAAGLPAGVLSVVGSGAIGDAFVEHPVPGLVTFGTSP 238
QY 234 VQORIGKLAG-----MRPIMLEGKDGAGIVLADADLDNNAKQIVAGAYVSGORCTAIK 289
Db 239 VGRNIGRIASGSEHLKHYALELGSNPFVVLGDADLEQVNAAPVGEKFLHQICIMALN 298
QY 290 VLVVEEVADELAEKISENVAKLSVGPDPDNATVTPVIDNSADPIESLVVDAROKAKE 348
Db 299 IIVEQPLLEDFTRRFRVERVKALPYGDPKPGVVPVINAQOLAGLKEKIKITAKAEGATL 358
QY 349 LNEFKRDRRLITPGLFDHVTLDMLAMEBPPGILPIIRVYDAEBAVAIAKNSDFGLSS 408
Db 359 LIGSGPQGNVMPVHFGVITATMEIARDEIFGPIVGLRLARDEAHLELANASEGLSSA 418

;; PRIOR APPLICATION NUMBER: US 60/094,190
 ;; PRIOR FILING DATE: 1998-07-27
 ;; NUMBER OF SEQ ID NOS: 33142
 ;; SEQ ID NO 27749
 ;; LENGTH: 594
 ;; TYPE: PRT
 ;; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-27749

Query Match 27.9%; Score 668; DB 4; Length 594;
 Best Local Similarity 33.2%; Pred. No. 8.2e-54;
 Matches 158; Conservative 90; Mismatches 212; Indels 16; Gaps 7;

QY 6 QNYVNGEKSSVN--QIEILSPIDSSLGFPVPMTRREVDHAKGREALPMAALTYVE 63
 DB 123 QAYVGDVADNDGQTIKVNNDPATGEIIGSVPMGAETRRALBAADKALPMAALTYVE 182
 QY 64 RAQYHKAADIIEERKEEIAATYLAKEISKAYNASTEVEVTRADLIRYAEBSIRLSTAD 123
 DB 183 RANKLRWFDMIEHQDLARLMTIEQKPLAEAKGEIAYAAFLPEWPEEKRIYDPTI 242
 QY 124 EGGKMDASTGHLAVIRQPVGIVLAAPYNPVNLGSKIAIPALIGVNVMEKPTQGS 183
 DB 243 FGHQPD-----KRIIVIKQPIGVTAITPMNPPSAMITRKAPALAGCTMWLKASQTP 297
 QY 194 VSGVLAKAFAPAGLPAAGVNTITGRSGEIGDYIYEHEEVNFINTGSPVQRIQK--L 241
 DB 298 YSALALAEIAPAGIPKGVFSVVTGSAGVGGELNSPIVRLKLTFTGSTEIGRLMAEGA 357
 QY 242 AGMRIMIELEGGKAGIVLADADLNNAAQIYAGAYDSGGQCTAIKRLVVEVADDELA 301
 DB 358 ODIKVSELELGNAPFIYFDADLDAVAEGALISKYRNNGQTCVAMRLYVODGVYDAFV 417
 QY 302 EKISENVAKLSVGDFPD--NATVTPVIDNSADPISLVVDAROKAKELNEFK---RDR 357
 DB 418 DKLKAQAAKLNIGNGLEAGCVTTGPLIDAKAVAKVEEHIDAVSKAKVSGSKPHALGCT 477
 QY 358 LITPGLFPHVTLDMLKAMEBPPPIIPITRVDAEBAVAIAANKSPFGLQSSVFTDFOKA 417
 DB 478 FPEPTLVDPKGNALVSKDETFGPIAPVFRFOBAEVLAMSDTEFGLASFYADDLARV 537
 QY 418 FDIANKLEVGTTHNNKTCR--GPDNPPFGLKSGAGVQIGIRYSIEAMTVKSIYL 472
 DB 538 FRVAQLEHGVGIN--TGLISNEVAPFGIGIKASGLGRGSKYGIEDYLEIKYLCL 591

RESULT 11
 US-08-976-063E-28
 ; Sequence 28, Application US/08976063E
 ; Patent No. 6524831
 ; GENERAL INFORMATION:
 ; APPLICANT: Steinhuchel, Alexander
 ; APPLICANT: Pfelebert, Horst
 ; APPLICANT: Rabenhorst, Uergen
 ; TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF CONIFERYL
 ; TITLE OF INVENTION: ALCOHOL, CONIFERYLALDEHYDE, FERULIC ACID, VANILLIN AND
 ; TITLE OF INVENTION: VANILLIC ACID AND THEIR USE
 ; FILE REFERENCE: Bayer-9998-CNO
 ; CURRENT APPLICATION NUMBER: US/08/976,063E
 ; CURRENT FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: 196 49 655.1 GERMANY
 ; PRIOR FILING DATE: 1996-11-29
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 28
 ; LENGTH: 481
 ; TYPE: PRT
 ; ORGANISM: not required under old rule
 US-08-976-063E-28

Query Match 27.9%; Score 667.5; DB 4; Length 481;
 Best Local Similarity 36.4%; Pred. No. 6.5e-54;
 Matches 169; Conservative 80; Mismatches 202; Indels 13; Gaps 6;

QY 15 SSVNOIEILSPIDSSLGFPVPMTRREVDHAKGREALPMAALTYVERAQLYHKAADI 74
 DB 16 SDETFEERSPLTGEVVSRAVAASLEDDADAAVAAGAFPMALAPSERARALLRADL 75
 QY 75 IERDKEEIAATYLAKEISKAYNASTEVEVTRADLIRYAEBSIRLSTADDEGKNDASTGH 134
 DB 76 LE-DRSSSEFTAAASFTGAAGWYGFNNVYLAAGMREAA-----AMTTOIQDVIIPSNVPG 129
 QY 135 KLAIVRRQPVGIVLAAPYNPVNLGSKIAIPALIGVNVMEKPTQGSGLVYAKAPA 194
 DB 130 SPMAVRQPCGVVIGIAFMNAPVILGVRAVMPALCGVTVLVKSSSELPFTHRLIGQVLH 189
 QY 195 BAGLPAGVNTITGRSGE---IGDYIYEHEEVNFINTGSPVQRIQKLAG--MRPYML 249
 DB 190 DAGIGDGVVNTIINAPQAPAVVERLIANPAVRNRNFGSTHYVGGIIGELSRHLKPAVL 249
 QY 250 ELGGKAGIVLADADLNNAAQIYAGAYDSGGQCTAIKRLVVEVADDELAKEISENVA 309
 DB 250 ELGGKAPPLVDDADLDAVAEAAAFGAIFYNOGOICMSTER-LIYVAVADAFVEKLAARKVA 308
 QY 310 KLSVGDPPDNMTVT--PVIDNSADPISLVVDAROKAKELNEFKRDGRLITPGLFDHYT 368
 DB 309 TIRAGDPPDPOSVGLSIDANAGORIOVLVDALAKGAROVVGGGLDSIMQPHLLDOYT 368
 QY 369 LDMKLAMEBPPPIIPITRVDAEBAVAIAANKSPFGLQSSVFTDFOKAPDIANKLEYGT 428
 DB 369 EEMRLYRESGPAVAVLRGPDDELLRLANDSEFGSAALFSRDSAMELAQVDSGI 428
 QY 429 VHNNKTCRGPDNPPFGLKSGAGVQIGIRYSIEAMTVKSIYL 472
 DB 429 CHINPPTHDEAOMPEFGVSKSGYGSFSGRASIEHFTQLRWLTI 472

RESULT 12
 US-09-538-092-1047
 ; Sequence 1047, Application US/09538092
 ; Patent No. 6753314
 ; GENERAL INFORMATION:
 ; APPLICANT: Glot, Loic
 ; APPLICANT: Mansfield, Traci A.
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
 ; FILE REFERENCE: 15966-542
 ; CURRENT APPLICATION NUMBER: US/09/538,092
 ; CURRENT FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: 60/127,352
 ; PRIOR FILING DATE: 1999-04-01
 ; PRIOR APPLICATION NUMBER: 60/178,965
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 1387
 ; SOFTWARE: CuraPatSeqFormatter Version 0.9
 ; SEQ ID NO 1047
 ; LENGTH: 517
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (0)-(0)
 ; OTHER INFORMATION: Polypeptide Accession Number P30837
 US-09-538-092-1047

Query Match 27.6%; Score 660; DB 4; Length 517;
 Best Local Similarity 34.8%; Pred. No. 3.7e-53;
 Matches 170; Conservative 80; Mismatches 201; Indels 38; Gaps 11;

QY 6 QNYVNGEKSSVNQ--IEILSPIDSSLGFPVPMTRREVDHAKGREAL--PMAALTY 60
 DB 38 QLFINNEQDQAVSKTPTNPTGVEVIGHVAEGRADVDRAVAARAFRLGSPWRMD 97
 QY 61 YVERAQLYHKAADIIEERKEEIAATY----LAKESKAYNASTEVEVTRADLIRYAEBSI 116
 DB 98 ASERGRLLNLADLVERDRVYLALETLDNGKPPQESYALDLDEVTK---VYRYFA---- 150


```
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Benson, Bob
/ REGISTRATION NUMBER: 30,446
/ REFERENCE/DOCKET NUMBER: PB248PP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 5241:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 508 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-781-986A-5241

Query Match      27.6%; Score 659.5; DB 4; Length 508;
Best Local Similarity 34.0%; Pred. No. 4e-53;
Matches 166; Conservative 99; Mismatches 198; Indels 25; Gaps 10;

QY 1 LTKEQNYNGEM--KSSVNOIEILSPIDSSLGFPVPMTRBEVDHAKAGREALPAMA 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 22 IAEITGLFNGEFVKSSDEETLEVTNPATGETLSHTTRAKDQVDPAVVAQEAESWLSL 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 LTVERRAOYLHKAADIIEERKEEIAVLAKETISKAVNASVTEVVTADLIRYAAEE-GIR 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 82 TSKSRRAQWLRIQGLKMAQKQKIMIEFLNNGKPIREF-----TADIPRAHHFHF 135
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 LSTSDGEGKMDASTGHKLAIVTRROPVGIYLAIAFYNYPNVLSGSKIAPALIGSNVMEK 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 136 ASVIEETEGTVNDIDKDTMSIVRHEBIGVVGAVVAMNPFMLAAMKIAIPALIAAGNTYIVIQ 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 PTTQGSVSGVLAKAFPAEAGLPAGVFNTITGRGSEIGDIYVEHEVNFINTGSTPVGQR 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 196 PSSSTPLSLLEVAKLFOEV-LPKGVNNTITGSGESGNALFNHDDGVDKLSFGSTDVGQ 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 238 IGLKLAG--MRPYMLEIGKGAGIYVLADADLDNAKOIVAGAYDSGORCTAIKRYLVVEE 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 255 VAAEAAKHLVPATLELGGKSANIIIDDAIIDLAVESIQGLIFNGEGVSSAGSRLLVHEK 314
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 VADDELAETISENVAKLSVGDPRDNAT-VTPVIDNSADFTESLVVDAROKGAKEL--NE 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 315 IYDQLVPRLOEAFSNIKVGNPDDEATQMGSGTGKQQLDKIQSYIDAAKESDAQIILAGHR 374
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 352 FRDGRLLTPGLFDVHTL-----DMKLAEEPPFGIILPIIVKAAEEVAIALNKSDREL 405
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 375 LTENG-LDKGFEEPTLIAVDNHHKLAOEIIFGVLTIVIKVDQOEALIDANDSEYGL 432
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 406 OSSVETRDPOKAFDIANKLEVGTVHINNKTGRGPDNFPPLGLKSGAGVQGIYSIEAMT 465
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 433 AGGVSSQNTLRALNTAKAVTRGRIMINT-YNGVPBGAFFPGYKKSIGETKKGALSNQ 491
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 466 NVKSLIVLD 473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 492 QVKNYIYID 499
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-09-328-352-4767
/ Sequence 4767, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 4767
/ LENGTH: 486
```

```
/ TYPE: PRT
/ ORGANISM: Acinetobacter baumannii
/ US-09-328-352-4767

Query Match      27.5%; Score 658; DB 4; Length 486;
Best Local Similarity 32.2%; Pred. No. 5.2e-53;
Matches 154; Conservative 92; Mismatches 219; Indels 14; Gaps 6;

QY 4 EYQNVYNGEMKSSVNO--IEILSPIDSSLGFPVPMTRBEVDHAKAGREALPAAALTV 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 14 QOQAYINQOMLAQSNATVPVSNPATGGEITGTPMGALEATQVAEALYTLQSKALTA 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 YERAOYLHKAADIIEERKEEIAVLAKETISKAVNASVTEVVTADLIRYAAEEGIRLSTS 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 74 QNRADIIILAMHKLVLIDHDELALIMTIEQKPLAEAKGEVRYAASFQWFAEEGRKI--- 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 ADEGKMDASTGHKLAIVTRROPVGIYLAIAFYNYPNVLSGSKIAPALIGSNVMEKPTQ 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 131 --YGDVIFTVNNQORFIIISKEPVGVAITPMNPFAMITRKAPALAAAGCTVYIKPANE 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 GSVSGVLAKAFPAEAGLPAGVFNTITGRGSEIGDIYVEHEVNFINTGSTPVGQRIGKL 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 189 TPYCALAIKLAERKAGIYAGVINYVTGKSQIEGVSFTSHKVKLTFTGSTIPVGRILMQ 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 AG--MRPYMLEIGKGAGIYVLADADLDNAKOIVAGAYDSGORCTAIKRYLVVEEVADE 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 249 CSSTIKLALIELGNAPLIFPDADLDKAVOGAIFAKFRNAGQCVCANRIYVHDNIYQA 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 LAEKISENVAKLSVGDPR--DNAVTVPVIDNSADFTESLVVDAROKGAKELNEPFRDG-- 356
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 309 FAEKRVQEVQKPKVGNGLGDEVOIGPLINERAVILKAQQLIDDAVSKGAKIACGGKQHALG 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 -RLTPGLFDVHTLDMKLAEEPPFGIILPIIVKDAEBAVALANKSDGLQSSVETRDPO 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 369 QTFPEPSVLTIVDRMEIVQEEIFGPVAPLRITFDEADVAVQAQNDITIGLAAYYSEIIS 428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 416 KAFDIANKLEVGTVHINNKTGRGPDNFPPLGLKSGAGVQGIYSIEAMTVKSLIVLDM 474
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 429 RLMRVSEQLLEVGMGM-NATAISNEVVPFGVKGSGVREGSKYGLSEFMTIKYWCGL 486
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 16
US-09-538-092-831
/ Sequence 831, Application US/09538092
/ Patent No. 6753314
/ GENERAL INFORMATION:
/ APPLICANT: Glot, Loic
/ APPLICANT: Mansfield, Traci A.
/ TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
/ FILE REFERENCE: 15966-542
/ CURRENT APPLICATION NUMBER: US/09/538,092
/ CURRENT FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/127,352
/ PRIOR FILING DATE: 1999-04-01
/ PRIOR APPLICATION NUMBER: 60/178,965
/ PRIOR FILING DATE: 2000-02-01
/ NUMBER OF SEQ ID NOS: 1387
/ SOFTWARE: CuratPatseqFormatter Version 0.9
/ SEQ ID NO 831
/ LENGTH: 500
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: Polypeptide Accession Number P00352
/ US-09-538-092-831

Query Match      27.5%; Score 658; DB 4; Length 500;
Best Local Similarity 34.1%; Pred. No. 5.4e-53;
Matches 163; Conservative 85; Mismatches 210; Indels 20; Gaps 8;

QY 8 YVNGEMKSSVN--QIEILSPIDSSLGFPVPMTRBEVDHAKAGREAL--PAAALTVY 62
```

```
Db      23 FINNEHDSVSGKKFVFNPNATEEELQVEEGKEVDKAAARQAFQSGWRMTDAS 82
Qy      63 ERAQVYHKAADIIEERDEKEIATVLAKEISKAY--NASVTEVTRADLIRYAEGBIRLSTG 121
Db      83 ERGRLLYKADLIERDLRLATMESNMGGKLYSNAYNLADLACTIKTLRYCAGWADKIQ-- 140
Qy      122 ADSGCKMDASTGKIAVIRQPIGIVLAIPYVYVNLSSKTAAPALLIGNVMFPRPQ 181
Db      141 ---GRTIPIDGNFPTTYTRHEPIGVCQIIPMNPPLMMLMKGPALSCNTVAVVPAEQ 196
Qy      182 GSVGVLAKAFPAEAGLPAGVFNTITGRGSEIDYVEHEEVNFINTGSPVGORIGKL 241
Db      197 TPLTALHVASLIEAGFPPEVNAVIVPEGYPTAAGALLSHMDIKNAFTSTEGKILKEA 256
Qy      242 AG---MRPIMLELGGKAGIVLADADLDNNAKOIVAGAYDYSQORCTAIRVLYVEEVD 298
Db      257 AGKSNLKRVTLELGGKSPCIVLADADLDNAVEFAHHGVFHQGCCIAASRIIVEESIYD 316
Qy      299 ELAEKISENVAKLSVGDPE--DNATVTPVIDNSADPIESLVDAKQKAK--ELMEFKR 354
Db      317 EFVRRSEVERAKKYLINPPLPGVTOGPOIDKEQYDILDLIESGKKEGATLECGGPGWGN 376
Qy      355 DGRLLTPGJEDHYTLDMKLAMEBPPGPIILPIRVKDAEBAVALANKSDPGLSSVFTPRD 414
Db      377 KGYFVOPFVSNVDEIRLAKKEIFGVQOIMFKSLDDVTKRANNTFYGLSGVFTKDI 436
Qy      415 OKAFDIANKLEVGTVHINNKTGRGPDNFPFLGLKSGAGVQGIYSIEAMTVKSYVL 472
Db      437 DKALITSSALQACTVAV--NCGYVVASQCPFGCFKMGNGELGEYGHETVEKTYTV 493
```

```
RESULT 17
US-09-602-777A-182
; Sequence 182, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krieger, Burkhard
; APPLICANT: Schroeder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602, 777A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932928.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
```

```
; PRIOR APPLICATION NUMBER: DE 19932933.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 182
; LENGTH: 497
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-09-602-777A-182
```

Query Match 27.2%; Score 650; DB 4; Length 497;

Best Local Similarity 33.3%; Pred. No. 36-52;

Matches 162; Conservative 86; Mismatches 201; Indels 38; Gaps 11;

```
Qy      8 YNNGEW--KSSVNOQIEILSPIDSSLSGFVPMTRREVDHAKRGREALPAMALTYEBA 65
Db      21 YINGSVVKAEGTQKRPVNDPAVGQEWGVSPEATASLDSVAGAKAFALSMASHLTAEET 80
Qy      66 OYLHKAADIIEERDEKEIATVLAKEISKAYNASVTEVTRADLIRYAEGBIRLSTADBG 125
Db      81 GYLKIKATIEBSRSEALNLTNRENGSPISETRGAASNAAGIFRYFA-----TLAPLDBE 136
Qy      126 GKMDASTGKIAVIRQPIGIVLAIPYVYVNLSSKTAAPALLIGNVMFPRPQSGVS 185
Db      137 DIRPFAGSAESIVDKDPDGVCALAPMNPPLNLVITKLAIPALLAGCTVIIRPASPTPLS 196
Qy      186 GLVLAKAFPAEAGLPAGVFNTITGRGSEIDYVEHEEVNFINTGSPVGORIGKLAA-- 243
Db      197 IFFITIEALIAAGVPPGVNLLTSGG--RFGDALVRHGVDKVAFPGSTPGKKIAAAGCEL 255
Qy      244 MRPIMLELGGKAGIVLADADLDNNAKOIVAGAYDYSQORCTAIRVLYVEEVDLAKR 303
Db      256 LRPVLTLELGGKSAIILPPADMSYLSRLIRSCMNTGTCTYISTRILAPSSRYAEVQGT 315
Qy      304 ISENVAKLSVGDPEFNATV--TPVIDNSADPIESLVDAKQKAK-----ELMEFKR 353
Db      316 VASTIAAGROQDPYDEETVFGPVASASQYSTWASYDSAREGARVAVAGTSSISISEGL 375
Qy      354 RGRLLTPGLPDHYTLDMKLAMEBPPGPIILPIRVYD--AAEVAIVANKSPFGLSSV 410
Db      376 ESSEGIQPTVPADVPPDMKISREELTGPVISTIKYDVTNGVEBALANNTKFGGLGLVF 435
Qy      411 TRDFKAFDIANKLEVGTVHIN---NKTGRGPDNFPFLGLKSGAGVQGIYSIEAMT-- 465
Db      436 GADDEGALLEVARQVDSGVGINFPGSNMBA-----PFGRRHESGNGVE---YGIKGLSA 486
Qy      466 --NVKSI 470
Db      487 YLTYSKI 493
```

RESULT 18

NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 9207
 LENGTH: 532
 TYPE: PRT
 ORGANISM: Human
 US-09-949-016-9207

Query Match 26.7%; Score 639; DB 4; Length 532;
 Best Local Similarity 33.7%; Pred. No. 3.6e-51;
 Matches 162; Conservative 81; Mismatches 212; Indels 26; Gaps 9;

```

QY 8 YVNGEMKSSVN--QIEILSPIDDSLGFPVPMTRREVDHAKAGREAL--PAMALTVY 62
DB 55 FNNEMQNSGKRVFVYVPATGEOYCEYQVADKADIDKAVQARLAFSLGSWMRMDAS 114
QY 63 ERAQYIHKADIIERDKKEIATVLA---KEISKAVNASVTEVPTADLIRYAAEGIRL 118
DB 115 EKGRLIDKLAIDVERRAVLATMESLNGSKPPLQAFYVLDQVTKT---FRYYAGWADKI 171
QY 119 STSADGKMDASTGHHKLAIVIRROPVGIYLAIPYVYVNLGSKTAPALIGNVMPKP 178
DB 172 H-----GNTIPVDGDYFTFTNHEPIGVCGQIIPMFPILMPAMKIAPLCCGNTVVIKP 225
QY 179 PTQGSVSGLVAKAFPAEAGLPAGVFNITIGRSEIGDIYVHEBVNFINTGSTPVGRI 238
DB 226 AEGTPLSALYMGALIEAGPPGVINILPGYGTAGALASHIGIDKIAFTGSTEVGKLI 285
QY 239 GKLAG--MRPIMLEIGKDAGIYVADADIDNAKQIVAGANDYSGORCTAIRVLYVEE 295
DB 286 QSAAGSNLKRVTLEIGSKSPNIFADADLDYAVEQAHQGVFPNQCCTAGSRIFVEES 345
QY 296 VADELAEKISENVAKLSVGDPPDNATVT--PVIDNSADPIESLVVADAROKAKELNEFK- 353
DB 346 IYEFVRBSVERAKRVVSSPPPTTEGQPDIDKKNYKILIELIGVABAKLECGGKG 405
QY 354 --RDGELLTPGLFDHVTLDMLKAMEBPGPIPLIRYKDAEBAVALANKSDPGLQSSVPT 411
DB 406 LGRKGFIEPTVFSNVTDMRIAKEEIRFGVQETILAFKTMDEVIBRANNSDGLVAAVPT 465
QY 412 RDPQKAFDIANKLEVTGVIHNNKTGSPDNFPFLGLKSGAGVQGIIRYSIEAMTVKSIY 471
DB 466 NQINKALTVSSAMQAGTVMINCYNALMAOS--PRGGPKMSGNEMGEFGIREVSEKVT 524
QY 472 L 472
DB 525 V 525

```

RESULT 23
 US-09-489-039A-7569
 Sequence 7569, Application US/09489039A
 Patent No. 6610836
 GENERAL INFORMATION:
 APPLICANT: Gaty Breton et. al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 TITLE OF INVENTION: PNEUMONIA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709 2004001
 CURRENT APPLICATION NUMBER: US/09/489,039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117,747
 PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 7569
 LENGTH: 498
 TYPE: PRT
 ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-7569

Query Match 26.5%; Score 635; DB 4; Length 498;
 Best Local Similarity 33.1%; Pred. No. 7.8e-51;
 Matches 161; Conservative 99; Mismatches 197; Indels 30; Gaps 11;

```

QY 1 LTKERYQNVNGEMKSSVNO--IEILSPIDDSLGFPVPMTRREVDHAKAGREALP--AW 56
DB 19 LAIEIRLPIINGEYCAADANTTFETIDPAAQQTLAGVAGKADVERAVKARQAQFNDNDW 78
QY 57 AALTYERQYIHKADIIERDKKEIATVLAKEISKAVNASV--TEVPTADLIRYAAEGIRL 115
DB 79 SQASPQRKALITRPADIMEAHEBELALLETLDTGKPIRHSIRDDIPAAAPAIRVYAE-- 136
QY 116 IRLTSADP--GGKMDASGHHKLAIVIRROPVGIYLAIPYVYVNLGSKTAPALIGNVV 174
DB 137 ----ALDKYGEVAPVTSNBLMTVREPIGVIAVWPMNPLIACMKLAGPALAAGNSV 191
QY 175 MEKPTQGSVSGLVAKAFPAEAGLPAGVFNITIGRSEIGDIYVHEBVNFINTGSTPV 234
DB 192 ILKPEKSPPLTALRLIAGLAKAGLPDGLVNVSGRGEHAGELALHPVEVITFTGSTR 251
QY 235 GORIGLAG--MRPIMLEIGKDAGIYVADA--DIDNAKQIVAGANDYSGORCTAIRV 290
DB 252 GKQLIKDAGDSNMKRVLEAGSKSANIVFADCPDQAVRATAGGI FYNQGVCTIAGRRL 311
QY 291 LVVEVADELAEKISENVAKLSVGDPPD--NATVTPVIDNSADPIESLVVADAROKAKEL 349
DB 312 LIEESIADEPLARLAKAQQHOPGNPDPDITTMGMLIDNTHADNVHST----RGSSQ 366
QY 350 NEFKSDGR-----LITPGLFDHVTLDMLKAMEBPGPIPLIRYKDAEBAVALANKSDP 403
DB 367 STLFLDKKNPMPAVGFTIFVDVDPASTLRBEIFGVVLVYTRKSEBELKLANDSDY 426
QY 404 GLOSSFTRDPQKAFDIANKLEVTGVIHNNKTGSPDNFPFLGLKSGAGVQGIIRYSIEA 463
DB 427 GIGAAVMTTRDLRAHMRSRRLKAGSVFVNNYND--GDMTVPPGQYKQSGNGRDKSLALEK 485
QY 464 MTNVKSI 470
DB 486 FTELKTI 492

```

RESULT 24
 US-09-543-681A-7151
 Sequence 7151, Application US/09543681A
 Patent No. 6605709
 GENERAL INFORMATION:
 APPLICANT: GARY BRETON
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709 1002-001
 CURRENT APPLICATION NUMBER: US/09/543,681A
 CURRENT FILING DATE: 2000-04-05
 PRIOR APPLICATION NUMBER: US 60/128,706
 PRIOR FILING DATE: 1999-04-09
 NUMBER OF SEQ ID NOS: 8344
 SEQ ID NO 7151
 LENGTH: 505
 TYPE: PRT
 ORGANISM: Proteus mirabilis
 US-09-543-681A-7151

Query Match 26.4%; Score 632.5; DB 4; Length 505;
 Best Local Similarity 32.7%; Pred. No. 1.4e-50;
 Matches 159; Conservative 97; Mismatches 209; Indels 21; Gaps 10;

```

QY 3 KEYQNVNGEMKSSVNO--IEILSPIDDSLGFPVPMTRREVDHAKAGREALPMAALT 60
DB 27 ESKYKILIGEWVSTSQOTTKYSPATGELLSEYSSGAEVDVLAVKARAKFDMRKTIS 86
QY 61 YERQYIHKADIIERDKKEIATVLAKEISKAVNASVT--EVPTADLIRYAAEGIRLS 119
DB 87 PAQRQALLKTIADRLGEQERPAWIESLDNGKPLNESKNIDLPASIDHFRYFA--GV-IR 143
QY 120 TSADGKMDASTGHHKLAIVIRROPVGIYLAIPYVYVNLGSKTAPALIGNVMPKP 179
DB 144 SHTDESAYLDENT--LSLVIREPIGVGOIIPMFPILMAAMKLAIPAIAGDVTVINPA 200

```



```
QY 119 STSADBGKMDASTGHLAV-----IRQPGIVLAIAPYVNVLSGSKIAPALIG 171
DB 138 -----GLVPALIEGOIPLRETSFYVTRREPLGVAGIAMVYPOVIAIMKSAAPLAAG 190
QY 172 NVVMFKPTGSGVGLAKAPABAGLPAGVFNITITGRGSEIGDIYVEHEENVINFTGS 231
DB 191 NAMIFRSEVTPPLTALKLAIITYEAGVDPDVFNVLISGREVOQMLTEHPLIKISTFG 250
QY 232 TPVGQRI--GKLGMRPIMLEIGKQDAGIVLADADLDNAKOIVAGAYDYSQORTAK 288
DB 251 TSGTKVMASASSSLKEVITMELGSKPLIIFPDADLDRAADIIVANAFSSQGVCTNGT 310
QY 289 RVLVVEVDELAEKISENVAKISVGDPPD-NATVTPVIDNSADFIESIIVDARQKAK 347
DB 311 RVPIHRSQARFPAKVLIERVQRIKGLDPDENTNFGPLVSPHMESVLGIYESGKQOKAR 370
QY 348 ELNEFKR-----DGLLTPGLFDHYTLDMKLAMEEPFGPIIPRIIRKDAEVAIAIK 400
DB 371 LLOGGERVTDGANGNAYVAPVFTDCSDMTIVREEIFGPMVSIILYDDEDEALIRAND 430
QY 401 SDFGLOSSEVTRDFOKAFDIANKLEVGTVHINNKTRGPDNFPPLGKSGAGVQIRYS 460
DB 431 TBYGLAAGVVTQDLARAHRAIHRLHLEAGICWINT-WGESPAEMPVGYGKSGVGRNGLLT 489
QY 461 IEAMTVKSIIVDM 474
DB 490 LAHYTRIKSVQVEL 503

RESULT 30
US-09-134-001C-4541
; Sequence 4541, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stramm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4541
; LENGTH: 510
; TYPE: PR1
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4541

Query Match 26.1%; Score 624; DB 3; Length 510;
Best Local Similarity 31.8%; Pred. No. 8.7e-50;
Matches 155; Conservative 101; Mismatches 209; Indels 22; Gaps 10;

QY 1 LTKXYQNVVNGEKSSVNV--QIEILSPIDSSLGFPVPMTRREVDHAMKAGREALPMAA 58
DB 23 IDBSYGLFINNERQASGSEETLVSNPNANGEDLAKAVABAKKQVDKAVQAHAHAPDSMSK 82
QY 59 LITYERAQYIHKADIIERPKERIATYLAKEISKAY-NASVTEVYRTADLIRYAAEBGIR 117
DB 83 ISKEERADYLLBISRRHEKTEHLATVESLONKPKYETSTIDVPOAANOFKYFASV-- 139
QY 118 LSTRSADBGKMDASTGHLAVIRROPVGIYLAIPYVNVLSGSKIAPALIGNVVMPK 177
DB 140 LITDEGSEVNEIDQNT--MSLVNNEPVGVGVYVAMNFPILLSKWLGRPLAAGNVVIO 196
QY 178 PPTQGSVSVGLVAKAPAEAGLPAGVFNITITGRGSEIGDIYVEHEENVINFTGSTVPQOR 237
DB 197 PSSSTPLSLIELAKIFQEV-LPKGVNVVLTKGSESGDAIFHHBGVDKLSFTGSTVDGYE 255
QY 238 IGLKLAGR--PIMLELGKQAGIVLADADLDNAKOIVAGAYDYSQORTAIRVLVVEE 295
```

```
DB 256 VAQAGAEIRVPTTLIELGSGSANIIFPDANLEQVIEBVGILFENQEVESAGSRLLVQSS 315
QY 296 VADELAEKISENVAKISVGDPPDNAT-VTPVIDNSADFIESTIVDARQ-----KGAK 347
DB 316 IYVELLPKLKEAFENIKVDDPFDEDTKMSAQCGPEQDLKIESYIKIAEBDDKANILITGGH 375
QY 348 ELNEFKRD-GRLLTPGLFPHVTLDMKLAMEEPFGPIIPRIIRYVDAEVAIAIKSPDLQ 406
DB 376 RTDNGLDKGYFEEPTIIEINDKHLQAOEELFGPVVVEKEDDEQEALEIANDSEYGLA 435
QY 407 SSVETRDFOKAFDIANKLEVGTVHINNKTRGPDNFPPLGKSGAGVQIRYSIAMTN 466
DB 436 GGIETTDIIRALNVAMAKMTGRIMWT-YNOIPAGAPFGGKYSKGI GREYVKDALKNYQO 494
QY 467 VKSIYLD 473
DB 495 VKNIFID 501

RESULT 31
US-09-328-352-6753
; Sequence 6753, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6753
; LENGTH: 493
; TYPE: PR1
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6753

Query Match 26.0%; Score 623; DB 4; Length 493;
Best Local Similarity 32.8%; Pred. No. 1e-49;
Matches 188; Conservative 94; Mismatches 214; Indels 16; Gaps 6;

QY 2 TKEYQN---YVNGEKSSVNVQ--IEILSPIDSSLGFPVPMTRREVDHAMKAGREALPMA 56
DB 8 TQPMQNVOLLINGQSVDAENQATFERISPIDGTVAATKAAATLEDVDRAIESQOAFKYM 67
QY 57 AALTYERAQYIHKADIIERPKERIATYLAKEISKAYNASVTEVYRTADLIRYAAEBGI 116
DB 68 SKLSPTERRLRILKADLDMDQKTEQPIETGMOETGSTATVGFNVHIAAMLEBEA---- 123
QY 117 RUSTSADBGKMDASTGHLAVIRROPVGIYLAIPYVNVLSGSKIAPALIGNVVMP 176
DB 124 -AMTQIDSSILPSDVPNGULAMGVPCGVIGIYIAWMNAPVILATRALMPLACGTVVL 182
QY 177 KPTQGSVSVGLVAKAPABAGLPAGVFNITITGRG---EIGDIYVEHEENVINFTGSTP 233
DB 183 KSEACPAATHRLRIEGLTEHAGLEGVGVNVYITAAEDAPQIVRLVSHRPVKINFTGSTK 242
QY 234 VQGRIGTKLGG--MRPIMLEIGKQDAGIVLADADLDNAKOIVAGAYDYSQORTAIRVYL 291
DB 243 VGIKIAETAKYUKPVLLELGGKAPVVLNEADINEAVNAVAFGAFNQGQICMSITERVL 302
QY 292 VVEEVADLEAEKISENVAKISVGDPPDNATVTPVID--NSADFIESTIVDARQAKELN 350
DB 303 VOONINADQPIETKLIETKRSIQAGNPESKONVIGVLESRAAURIQHLLEDAQVKGADLPL 362
QY 351 EFRGDSRLLTPGLFDHYTLDMKLAMEEPFGPIIPRIIRYVDAEVAIAIKSPDLQ 422
DB 363 GHIEEDTTPQPLVLNIKPDMLLYREESFGVCTVQRFSSIEGVALANDSEFGSSAVP 442
QY 411 TRDFOKAFDIANKLEVGTVHINNKTRGPDNFPPLGKSGAGVQIRYSIAMTVKSI 470
DB 423 SONISQALEVAQOISDGICHINGATVHDEAQNPFQGTSSGIGRFGSKYSVAEFTDLRWI 482
```

Qy 471 VL 472
 Db 483 TI 484

RESULT 32

US-09-651-941-9
 ; Sequence 9, Application US/09651941
 ; Patent No. 6355470

GENERAL INFORMATION:

APPLICANT: ROUVIER, PIERRE E
 APPLICANT: WALTERS, DANA M
 APPLICANT: RAINIER, RUSS
 TITLE OF INVENTION: Genes Encoding Picric Acid Degradation
 FILE REFERENCE: BC1022 US NA
 CURRENT APPLICATION NUMBER: US/09/651,941
 PRIOR FILING DATE: 2000-08-31
 PRIOR APPLICATION NUMBER: 60/152,545
 NUMBER OF SEQ ID NOS: 28
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 9
 LENGTH: 485
 TYPE: PRT
 ORGANISM: Rhodococcus erythropolis HL PM-1

US-09-651-941-9

Query Match 26.0%; Score 621; DB 3; Length 485;
 Best Local Similarity 34.0%; Pred. No. 1.5e-49;

Matches 156; Conservative 80; Mismatches 205; Indels 18; Gaps 7;

Qy 23 LSPIDSSLGVPYAMTREVVDHAKGREALPMAALTVYERAOYLHKAADIIRDKKEI 82
 Db 28 INPADGSHLASVAEATADVARAVERAKAARTQMRPAQRTRLMFRYAALIEHKTEL 87
 Qy 83 ATVLAKESKAYNASV-TEVVRTADLIRYAABEGRILSTSADEGKMDASTGHKLAIVRR 141
 Db 88 AOLQSRDMGKPIRESIGIDLPIMIETLEYFA---GLVTKIE--GRTTPAPGRFLNYTLR 141
 Qy 142 QPVGVILAIAPYNYVNLGSKIAAPALIGNVMPKPTQGSVSGVLAKAPAEAGLPAG 201
 Db 142 EPIGVGAILTPWNPFAVQAWKIAPLAMGNALIVKPAQLAPVPALELLEAGLPAG 201
 Qy 202 VENTTIGRGEISGIDYVEHEEVNFINFTGTPVQGRIGKLAGMRPI--MLEIGKDAIV 259
 Db 202 LVNVLPGRGSVAGNALVQHPVSKVTFGTSTEVGOQIGRMADRLITASLEIGKSAIYA 261
 Qy 260 LADADLDNAKOIVAGAYVYSGQRTAIRVLVVEEVADELAKESENVAKLSVGDPPDN 319
 Db 262 FGDSPPKAAVAAVFQAMYSNOGETCTAPSRLLVERPIYDEVVELVQARVEAARVGDPLDP 321
 Qy 320 AT-VTPVIDNSADPIESLVDAROKGAKEIN-----EPRKDRGLLTPGLFDHTLDM 371
 Db 322 DTEIGPLISAEORESVHVSSTEGATLISGGDQSPFGAPBQGYRPTLFSGVATDM 381
 Qy 372 KLAMEEPFPIIPIIRVDAEBAVALANKSDPGLQSSVFTPDFOKAFDIANKLEVGTYHI 431
 Db 382 RIAREEIFGPVLSVLPFEGEBEAITLANDTVFGLAAGVTRDVGRLRPAQTLIDAGNWI 441
 Qy 432 NNKTRGPDNPFPFLGLKSGAGVQGIKRSISAMTNVKS I 470
 Db 442 NSWGVLPAS-PYRGFGSGYSDLDGAAIESTTKKESI 479

RESULT 33

US-09-955-597-9
 ; Sequence 9, Application US/09955597
 ; Patent No. 6461856

GENERAL INFORMATION:

APPLICANT: ROUVIER, PIERRE E
 APPLICANT: WALTERS, DANA M
 APPLICANT: RAINIER, RUSS

FILE OF INVENTION: Genes Encoding Picric Acid Degradation

FILE REFERENCE: BC1022 US NA
 CURRENT APPLICATION NUMBER: US/09/955,597

CURRENT FILING DATE: 2001-09-17
 PRIOR APPLICATION NUMBER: 60/152,545

PRIOR FILING DATE: 1999-10-03
 NUMBER OF SEQ ID NOS: 28

SOFTWARE: Microsoft Office 97
 SEQ ID NO 9

LENGTH: 485
 TYPE: PRT

ORGANISM: Rhodococcus erythropolis HL PM-1

US-09-955-597-9

Query Match 26.0%; Score 621; DB 4; Length 485;
 Best Local Similarity 34.0%; Pred. No. 1.5e-49;

Matches 156; Conservative 80; Mismatches 205; Indels 18; Gaps 7;

Qy 23 LSPIDSSLGVPYAMTREVVDHAKGREALPMAALTVYERAOYLHKAADIIRDKKEI 82
 Db 28 INPADGSHLASVAEATADVARAVERAKAARTQMRPAQRTRLMFRYAALIEHKTEL 87
 Qy 83 ATVLAKESKAYNASV-TEVVRTADLIRYAABEGRILSTSADEGKMDASTGHKLAIVRR 141
 Db 88 AOLQSRDMGKPIRESIGIDLPIMIETLEYFA---GLVTKIE--GRTTPAPGRFLNYTLR 141
 Qy 142 QPVGVILAIAPYNYVNLGSKIAAPALIGNVMPKPTQGSVSGVLAKAPAEAGLPAG 201
 Db 142 EPIGVGAILTPWNPFAVQAWKIAPLAMGNALIVKPAQLAPVPALELLEAGLPAG 201
 Qy 202 VENTTIGRGEISGIDYVEHEEVNFINFTGTPVQGRIGKLAGMRPI--MLEIGKDAIV 259
 Db 202 LVNVLPGRGSVAGNALVQHPVSKVTFGTSTEVGOQIGRMADRLITASLEIGKSAIYA 261
 Qy 260 LADADLDNAKOIVAGAYVYSGQRTAIRVLVVEEVADELAKESENVAKLSVGDPPDN 319
 Db 262 FGDSPPKAAVAAVFQAMYSNOGETCTAPSRLLVERPIYDEVVELVQARVEAARVGDPLDP 321
 Qy 320 AT-VTPVIDNSADPIESLVDAROKGAKEIN-----EPRKDRGLLTPGLFDHTLDM 371
 Db 322 DTEIGPLISAEORESVHVSSTEGATLISGGDQSPFGAPBQGYRPTLFSGVATDM 381
 Qy 372 KLAMEEPFPIIPIIRVDAEBAVALANKSDPGLQSSVFTPDFOKAFDIANKLEVGTYHI 431
 Db 382 RIAREEIFGPVLSVLPFEGEBEAITLANDTVFGLAAGVTRDVGRLRPAQTLIDAGNWI 441
 Qy 432 NNKTRGPDNPFPFLGLKSGAGVQGIKRSISAMTNVKS I 470
 Db 442 NSWGVLPAS-PYRGFGSGYSDLDGAAIESTTKKESI 479

RESULT 34

US-09-655-270A-9
 ; Sequence 9, Application US/09655270A
 ; Patent No. 6329151

GENERAL INFORMATION:

APPLICANT: Rouviere, Pierre E.
 TITLE OF INVENTION: High Density Sampling of Differentially Expressed Prokaryotic MR
 FILE REFERENCE: BC1011 US NA
 CURRENT APPLICATION NUMBER: US/09/655,270A
 PRIOR FILING DATE: 2000-09-05
 PRIOR APPLICATION NUMBER: 60/120,702
 PRIOR FILING DATE: 1999-February-19
 PRIOR APPLICATION NUMBER: 60/152,542
 NUMBER OF SEQ ID NOS: 37
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 9

LENGTH: 508
 TYPE: PRT

ORGANISM: Rhodococcus erythropolis HL PM-1

US-09-655-270A-9

Query Match 26.0%; Score 621; DB 3; Length 508;
Best Local Similarity 34.0%; Pred. No. 1,7e-49;
Matches 156; Conservative 80; Mismatches 205; Indels 18; Gaps 7;
QY 23 LSHIDSSIGFVPMATREEDVHAMKAGREALTVYERAOYLHKADIIERDEET 82
DB 51 INPADSSHLSAVEATADVARAVEAKAAFTWQRNRPQRTLRMERVAALIEHKTEL 110
QY 83 ATYLAELISKAYNASV-TEVVRTADLIRYAEIGIRLSTSDGCKMDSTGHKLAVIR 141
DB 111 AQAQSDMGKPIRESIGIDLPIMIELEFYFA---GLVTKIE--GRTTPAPGRFLMYTJR 164
QY 142 OPVGIVLAIAPYNYPNVLGSKTAPALIGANVMFPPTQGSVGLVAKAFAPAGLPAG 201
DB 165 EPFGVGAITPMFPFPAVOAWKIAIPALMGNALVLPQAOLAPVPAALGELAEAGLPDG 224
QY 202 VFNITIRGSEIDYIYEHEEVNFINTGSTPVGQRIKLAGMRPI--MLELGGKAGIY 259
DB 225 LVNVLPGRGSVAVNALVQHPVSGKVTFTGSTEVGQOIGRAADRLITASIELGKSAVA 284
QY 260 LADADIDNAKQIVAGAYDYGQRTAKRVLYVEEVADELAEKISENVAKLSVGPDPN 319
DB 285 FCGSSPKAAVAAYVFOAMYSNQGETCTAPSRLIVERPIYDEVELVQARVEAAVGDPLP 344
QY 320 AT-VTFVIDNSADFIESLVVDAROKAKELN-----EFKRDGRLITPGLFDHTLDM 371
DB 345 DTEIGPLIAEORESHSVSYSGTEGATLISGGDQSPGAPRQGYRPTLFSGVTAOM 404
QY 372 KLAMEEPFGILPIIRVKDAEVAVALANKSDPGLQSSVFTRDFOKAFDIANKLEVGTVHI 431
DB 405 RIRKEIFPGVLSVLPFEGEEALITLANDVFLAAGVFLRDVGRRLRPQITIDAGNVI 464
QY 432 NNKTRGPDNPFPLGKSGAGVQGIKRSIEAMTNVKS1 470
DB 465 NSWGVLPAS-PYRGGSGSGYGDGLQAALIESFTKEKSI 502
RESULT 35
US-09-538-092-851
Sequence 851, Application US/09538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Glaxo, Inc.
APPLICANT: Menzies, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538, 092
CURRENT FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CuratSeqFormatter Version 0.9
SEQ ID NO 851 ;
LENGTH: 517 ;
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: Polypeptide Accession Number P05091
US-09-538-092-851
Query Match 26.0%; Score 621; DB 4; Length 517;
Best Local Similarity 33.7%; Pred. No. 1,7e-49;
Matches 162; Conservative 85; Mismatches 212; Indels 22; Gaps 10;
QY 6 QNVNGEWSVNO--IELISPIDSSLGFPVPMATREEDVHAMKAGREAL--PAWAALT 60
DB 38 QIFINNEWHDAVSRKTFPTVNPSTGEVLCQVAEGDKEDVDKAVKAAAFQSGSPRRMD 97
QY 61 VVERAOYLHKADIIERDEKELATVLAKEISKAYNAS-VTEVVRTADLIRYAEIGIRLS 119

DB 98 ASHRRGLNRLADLIERDSTYLALETUNGKRPYISVYVDMDMLKGRYYA----- 150
QY 120 TSADE-GGMDASTGHKLAVIRQVGIYLAIPYNYPNVLGSKTAPALIGANTVMPK 178
DB 151 GNADKXHGKTIPIIDGFFSYTRHEPVGCGQIIPNNFPLMQAMKLGPALATAGNVVMKV 210
QY 179 PQGVSGLVLAKEPAENLPGVFNITIRGSEIGDYIYEHEEVNFINTGSTPVGQRI 238
DB 211 ABQPLTALYVNLKEAGFPFGVNVIVFGFPTAGAAASHEDVDKAFSTSTELGRVI 270
QY 239 GKLAG---MRPIMLEIGKDAIVLADADIDNAKQIVAGAYDYGQRTAKRVLYVEE 295
DB 271 QVAAGSNKRTLELGKSPNIIIMSDMDVADEAHALFPNQCQCCASRTVQGD 330
QY 296 VADELAEKISENVAKLSVGPDPNATVT--PVIDNSADFIESLVVDAROKAKEL--NEF 352
DB 331 IYDEFERSVAKSRVGNPFDSTKEGQPVDEQFKILGYINTGQEGAKLLCGGSI 390
QY 353 KRD-GRLLTPGLFDHYTLDMKLAMEEPFGILPIIRVKDAEVAVALANKSDPGLQSSVFT 411
DB 391 AADRGYFIOPTVFGVODGMTIAKEIFGPVQIILFKTIEBVVGGRANNSTYGLAAVPT 450
QY 412 RDPKAFDIANKLEVGTVHIINNKTGRGPDNPFPLGKSGAGVQGIKRSIEAMTNVKSIV 471
DB 451 KOLDKRNLYSQLQGTWVNCYDVFGAOS-PFGGIKSGSGRELGEGLOHYTEVKIVT 509
QY 472 L 472
DB 510 V 510
RESULT 36
US-09-949-016-10653
Sequence 10653, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYOMPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 10653
LENGTH: 520
TYPE: PRT
ORGANISM: Human
US-09-949-016-10653
Query Match 26.0%; Score 621; DB 4; Length 520;
Best Local Similarity 33.7%; Pred. No. 1,7e-49;
Matches 162; Conservative 85; Mismatches 212; Indels 22; Gaps 10;
QY 6 QNVNGEWSVNO--IBLISPIDSSLGFPVPMATREEDVHAMKAGREAL--PAWAALT 60
DB 41 QIFINNEWHDAVSRKTFPTVNPSTGEVLCQVAEGDKEDVDKAVKAAAFQSGSPRRMD 100
QY 61 VVERAOYLHKADIIERDEKELATVLAKEISKAYNAS-VTEVVRTADLIRYAEIGIRLS 119
DB 101 ASHRRGLNRLADLIERDRTYLALETUNGKRPYISVYVDMDMLKGRYYA----- 153
QY 120 TSADE-GGMDASTGHKLAVIRQVGIYLAIPYNYPNVLGSKTAPALIGANTVMPK 178
DB 154 GNADKXHGKTIPIIDGFFSYTRHEPVGCGQIIPNNFPLMQAMKLGPALATAGNVVMKV 213

QY 179 PTQSVSGVLAKAFAGAGLPGAVFNTITGRGSEIGDYIHEHEVNFINTGSTPVGQRI 238
 DB 214 AQGTPLTALYVNLIKKAGFPFGVNVIVBGFPTAGALASHEDVDKVAFTSGTEIGRVI 273
 QY 239 GKLAG---MRPIMLEIGGDAQIVLADLDLNAKQIVAGAVDYSQRCRTAIKRVLYVEE 295
 DB 274 QVAAGSSNKRVTLELGGKSPNIMSADMDMAVEQHPALFENOQCCACASRTFVQED 333
 QY 296 VADELAEKISENVAKLSVGDPPDNATVT--PVIDNSADFIISLVYDARQKAKEL--NEF 352
 DB 334 IYDEFERSVARAKSVVGNPFDSTEQGPQVDETQFKILGINTGQEGAKLGGGI 393
 QY 353 KED-GRLLTPGLFDHVTLLDMKLAMEBPPGPIILIRVKDAEBAVALANKSDGLOSVFT 411
 DB 394 AADRGGFIQPTVEGVDQDGMTAKEEIEFGVWQILFKTIEVYGRANNSTYGLAAAVPT 453
 QY 412 RDPQAFDIANKLEVTVTTHINKTGPGPNPPLGLKSGAGVQGRISYEAMTNYKSIIV 471
 DB 454 KDLDRANYSLOALQAGTVMWNCYDVGAQS-PFGYKMSGSGRELGELQAYTEVKVT 512
 QY 472 L 472
 DB 513 V 513

RESULT 37
 US-09-328-352-6687
 ; Sequence 6687, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 6687
 ; LENGTH: 502
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-6687

Query Match 25.8%; Score 618; DB 4; Length 502;
 Best Local Similarity 31.3%; Pred. No. 3,1e-49;
 Matches 149; Conservative 97; Mismatches 206; Indels 24; Gaps 7;

QY 6 QYVNGEMKSSVNO--TEILSPIDSSLGFPVPMTRREVDHANKAGREALPMAALTYVE 63
 DB 29 QAFIAGKWDADHQOTSILNPATLEIGTVPWNGRAEAEALAEAKEMPLMKNTAKD 88
 QY 64 RAQYLHKADIIERDKKEIATVLAKEISKAYNASVEVVRTADLIRYABEGIRLSTSD 123
 DB 89 RSILKKMPDLIISNDELAFILTSBQKPLAEKELIYASFIEMFAEAKRYVGDII 148
 QY 124 EGGKMDASTGSHKLAIVRRQVGIIVLAIPVNYVNSGSKIAPALIGVNVMPKPTQGS 183
 DB 149 BPPYPRAR-----IYVKNQPIGVVAITTPNFPAMITRKVAPALAAQPCIVKAPETP 203
 QY 184 VSGVLAKAFAGAGLPGAVFNTITGRGSEIGDYIHEHEVNFINTGSTPVGQRIKLAG 243
 DB 204 FALALVLDAVQAGVPAIFSVITGDVHIGDAIFESDVVRKFTFGSTPVGMLLERSA 263
 QY 244 --MRPIMLEIGGDAQIVLADLDLNAKQIVAGAVDYSQRCRTAIKRVLYVEEVADELA 301
 DB 264 KTLKVSLSLEGGAPPIVFDADDAIIGALIAKRRNAGQTCVCNRFVLQAGIYEKPI 323
 QY 302 EKISENVAKLSVGDPPD-NATVTVIDNSADFIISLVYDARQKAKEL--NEFQRDR 357
 DB 324 AALSOQIKQNFNIGNGLEAGHDIGLPLINANAIVKVEAHIDALDKNRLVVGKKGKAGEL 383
 QY 358 LITPGLFDHVTLLDMKLAMEBPPGPIILIRVKDAEBAVALANKSDGLOSVFTRFOKA 417

DB 384 FFEPTLIADVDAMDVAQTQETFGPLAAVKEETEQOAVEMANATEFGLAAYCYTKOLGRA 443
 QY 418 FDIANKLEVTGTHIN-----NKTGRGPNPPLGLKSGAGVQGRISYEAMTNYK 468
 DB 444 WRMSQLEIRGVNIGVINKGLISNEVA-----PFGGIKQSGLGRSGSKYGIEDYLEIK 493

RESULT 38
 US-09-949-016-9559
 ; Sequence 9559, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CLO01307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9559
 ; LENGTH: 439
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-949-016-9559

Query Match 25.8%; Score 616.5; DB 4; Length 439;
 Best Local Similarity 34.4%; Pred. No. 3.5e-49;
 Matches 151; Conservative 79; Mismatches 184; Indels 25; Gaps 8;

QY 8 YVNGEMKSSVNO--QIEILSPIDSSLGFPVPMTRREVDHANKAGREAL--PMAALTYV 62
 DB 2 FINNEMHSSKSGKFKATCNPSTREQICEVEBGDKPDVDKAVEAQAQVORSGSPWRRLDAL 61
 QY 63 ERAQYLHKADIIERDKKEIATVLAKEISK---AYNASVTEVVRTADLIRYABEGIRL 118
 DB 62 SGRGLHQLADLYERDRATLALLETMDGKPELHAFIDLEGCIPT--LRYFAGMADKI 118
 QY 119 STSADGGMDASTGSHKLAIVRRQVGIIVLAIPVNYVNSGSKIAPALIGVNVMPKPT 178
 DB 119 Q-----GKTIPTDNNVCFTHHEPIGVCAITTPNFPPLMLWMKLAPALCCGNTMVLKP 172
 QY 179 PTQSVSGVLAKAFAGAGLPGAVFNTITGRGSEIGDYIHEHEVNFINTGSTPVGQRI 238
 DB 173 AQGTPLTALYVNLIKKAGFPFGVNVIVBGFPTVGAALASHHPQINKIAFTGSTEVGKLV 232
 QY 239 GKLAG---GMRPIMLEIGGDAQIVLADLDLNAKQIVAGAVDYSQRCRTAIKRVLYVEE 295
 DB 233 KEAAGRSMLKRVTLLELGGKNPCIVCADADLDLAVECAHQGFENQGCCCTAASRVFVEQ 292
 QY 296 VADELAEKISENVAKLSVGDPPDNATVT--PVIDNSADFIISLVYDARQKAKEL--ELNE 351
 DB 293 VTSSEFVRSVVEYAKRPVGDPPDVTQEQPQDQKQFKILIELLSGKKEGAKLECGGSA 352
 QY 352 FKRDRLTPGLFDHVTLLDMKLAMEBPPGPIILIRVKDAEBAVALANKSDGLOSVFT 411
 DB 353 MEDKGLFIKPIYFSVTONMRAKEIEIGPVPILKFSIEBVIKRAINSTDVGLTAAVFT 412
 QY 412 RDPQAFDIANKLEVTGTH 430
 DB 413 KDLDRALKLASALIESGTIV 431

RESULT 39
 US-09-155-183-4
 ; Sequence 4, Application US/09155183
 ; Patent No. 6323011


```

QY 3 KEYQYVNGEMKSVNQIIEILSPIDSSLGFPVPMATREEVDAHAKGREALP--AWAALT 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 14 REFGEFFIDEMWAKGDFDRSSPAHDVPVTRIPRCTREDLDEAVAAARAENGSMAGIA 73
QY 61 VYERQYVHLKADIIERDKERIATVLAKEISKAYNASTEVEVTRADLIRYAABEGIRLST 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 AADRAAVLLKKAAGLIERRDDIAWEVLENGKRIISQAKGIDHCTACFEMAAGAARMLH- 132
QY 121 SADEGKMDASTGHL-AVIRROPVGIYLAIPYNPVNLGSKIAAPALIGNVVMFKXP 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 -----GDTFNNIGEGLFQGNVLEPIVVGILTPMNPFMILCERAPFIILASGCTLVKPA 187
QY 180 TQGSVGLVLAFAFAEAGLPAGVNTITGRGSEIGDYVEHEBVNFINTGSTPVGORIG 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 EVTSATVLLAEIILADAGLPKGVFNVTGTGRVQGMTEHQDIDMLSTGSTGVGKSCI 247
QY 240 KLA---GMRPIMLEGGKADAGIVLADADLDAAKQIVAGAYDSGRCPTAIRVLVEEV 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 HAAADSNIKKLGLEIGKNPIVFADSNLEDAADAVAFISFNTGCCVSSRLIVERSV 307
QY 297 ADELAKISENVAKLSVDPFDNAT-VTPVIDDNSADFIESIYVDAROKAKEL-----N 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 AKKPERLVVPKMEKIRKVDPPDPETQIGAITTEAQNKTLLDYIAKKAEGAKLLCGGGIV 367
QY 351 EFKRDGRLITPGLFDHVTLLDMKLAMEBPPGPIILPIRVKDAEBAVANKSDFGLOSVF 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 DFGK-GQYIQPTLFTDVKPSMGIADEIFGPVLASFHPTVDEALAINADTYGGLAAVW 426
QY 411 TRDPKADPIANKLEVGTVHINKTGRGPDNPFPLGLKSGAGVCGIRISIEAMTVKSI 470
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 427 SKDIDKALAVTRVRAGRFWNTTWSGGPBT-PLGGFKOSGWRGGLVGEVEYQIKSV 485
QY 471 VLD 473
Db 486 HIE 488

```

```

RESULT 45
US-09-118-317-2
; Sequence 2, Application US/09118317
; Patent No. 6197562
; GENERAL INFORMATION:
; APPLICANT: Niwa, Mineo
; APPLICANT: Saito, Yoshimasa
; APPLICANT: Ishii, Yoshiro
; APPLICANT: Yoshida, Masaru
; APPLICANT: Suzuki, Hiromi
; TITLE OF INVENTION: No. 6197562el L-borbose Dehydrogenase and No. 6197562el
; TITLE OF INVENTION: L-borbose Dehydrogenase Obtained from Gluconobacter
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Obion, Splyak, McClelland, Maier & Neustadt, P.C.
; STREET: 1755 Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS-DOS Editor
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/118,317
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/513,841
; FILING DATE: 01-NOV-1995
; APPLICATION NUMBER: UK 9304700.9
; FILING DATE: 08-MAR-1993
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: JP 241851/1993
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: NORMAN F. OBLON
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-909-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 497 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Gluconobacter oxydans
; STRAIN: T-100
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1..497
; IDENTIFICATION METHOD: experimentally
US-09-118-317-2

```

```

Query Match      25.4%; Score 609; DB 3; Length 497;
Best Local Similarity 30.8%; Pred. No. 2,1e-48;
Matches 149; Conservative 101; Mismatches 213; Indels 20; Gaps 8;

QY 3 KEYQYVNGEMKSVNQIIEILSPIDSSLGFPVPMATREEVDAHAKGREALP--AWAALT 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 14 REFGEFFIDEMWAKGDFDRSSPAHDVPVTRIPRCTREDLDEAVAAARAENGSMAGIA 73
QY 61 VYERQYVHLKADIIERDKERIATVLAKEISKAYNASTEVEVTRADLIRYAABEGIRLST 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 AADRAAVLLKKAAGLIERRDDIAWEVLENGKRIISQAKGIDHCTACFEMAAGAARMLH- 132
QY 121 SADEGKMDASTGHL-AVIRROPVGIYLAIPYNPVNLGSKIAAPALIGNVVMFKXP 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 -----GDTFNNIGEGLFQGNVLEPIVVGILTPMNPFMILCERAPFIILASGCTLVKPA 187
QY 180 TQGSVGLVLAFAFAEAGLPAGVNTITGRGSEIGDYVEHEBVNFINTGSTPVGORIG 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 EVTSATVLLAEIILADAGLPKGVFNVTGTGRVQGMTEHQDIDMLSTGSTGVGKSCI 247
QY 240 KLA---GMRPIMLEGGKADAGIVLADADLDAAKQIVAGAYDSGRCPTAIRVLVEEV 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 HAAADSNIKKLGLEIGKNPIVFADSNLEDAADAVAFISFNTGCCVSSRLIVERSV 307
QY 297 ADELAKISENVAKLSVDPFDNAT-VTPVIDDNSADFIESIYVDAROKAKEL-----N 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 AKKPERLVVPKMEKIRKVDPPDPETQIGAITTEAQNKTLLDYIAKKAEGAKLLCGGGIV 367
QY 351 EFKRDGRLITPGLFDHVTLLDMKLAMEBPPGPIILPIRVKDAEBAVANKSDFGLOSVF 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 DFGK-GQYIQPTLFTDVKPSMGIADEIFGPVLASFHPTVDEALAINADTYGGLAAVW 426
QY 411 TRDPKADPIANKLEVGTVHINKTGRGPDNPFPLGLKSGAGVCGIRISIEAMTVKSI 470
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 427 SKDIDKALAVTRVRAGRFWNTTWSGGPBT-PLGGFKOSGWRGGLVGEVEYQIKSV 485
QY 471 VLD 473
Db 486 HIE 488

```

```

RESULT 46
US-09-134-001C-4451
; Sequence 4451, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

```

PRIOR APPLICATION NUMBER: US 60/090,717
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.2
SEQ ID NO: 22
LENGTH: 501
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-344-862-22

Query Match 25.4%; Score 607; DB 4; Length 501;
Best Local Similarity 31.0%; Pred. No. 3.3e-48;
Matches 151; Conservative 98; Mismatches 204; Indels 34; Gaps 9;

QY 8 YNCEW--KSSVNOIEILSPIDSSLGFPVPMTRVEVDHAKRGREALP--AMALTYVE 63
DB :
23 FINGQFIDAASGKTETTPRNGEVLATIAEGDKEDVDLAVNAARYAFDHPMPRTMGE 82
QY 64 RAQYLHKADIIERKEEITATLAKESKAYN-ASTEVEVRTADLIYYAEBSGRISTSA 122
DB :
83 RAKLINRKADLLIENIEBELAKLDADVGGKLFQGLKYADIPATGHPRYN----- 132
QY 123 DEGRMDASTGHKLAVIR-----ROPVGIVLAIPYNYPVNLSGSKTAPALIGNVV 174
DB 133 ---GAADKIHGELLMTROSLEFGYTLAKEPIGVGNIIIPNNPSIMPATKYAPMAAGCTM 189
QY 175 MEKPPTQSVSGLVLAKAPAAAGLPAGVENITTGRCSEIGDYIVEHEEVNFIFTSSTPV 234
DB 190 VVKPAEQSLSLPAHLHSKEAGIPDGVTINVTGFCSTGAALASHMDVDKVSFTGSTDV 249
QY 235 GQRISKLA---GMRPIMFLGKGDAGIVLADADDLNAKOIVAGANDYSQRCITATRYL 291
DB 250 GRKITQAAAASNLIKCVSIELGKKSPLLIENDADIDRAADLALLGCFFNKGEICVASRVF 309
QY 292 VVEHVADLAEKISSENVAKLSTVGDPFDN-ATVPVTDNSADPTIESLVVAROKAKELN 350
DB 310 VOEGYDVKEVELKAKDMWTGVPDPSSTARQQPYQVDKROFEKILTSTIEHCKNEGATLLT 369
QY 351 EFKR---DGRLLTPGLFDHVTLDMLKLAWEPPGPILPIIRVKDAEEVAVALANSDFGIQS 407
DB 370 GSKAIGDGYFIQPIPIPADVTEDMKIYDEIGEPVWSLMKKFIVEBGIKCANNTKYGLAA 429
QY 408 SVTFRDPOKARDIANKEVGTVINNNKTRGCPDNPFPLGLKSGAGVQGIRISLEANTVN 467
DB 430 GILSQDIDLINTVSRISKAGIIWNVCYGFDD-DPYGGYKMSGNGRESGDALDNYLOT 488
QY 468 KSIYLDM 474
DB :
489 KSVVMP 495

RESULT 48
US-09-328-352-5847
Sequence 5847, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 5847
LENGTH: 498
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-5847

Query Match 25.3%; Score 604.5; DB 4; Length 498;
Best Local Similarity 33.3%; Pred. No. 5.7e-48;
Matches 160; Conservative 85; Mismatches 215; Indels 21; Gaps 9;


```
Db      337  ORKIDDPLEEDRMRPBLINRPHLEVLGVPVKAAGKAVLGGDIYVEDEPKLKDGY 396
      359  LTFGLFDHVTLDNKLAMEEPFGILPIIRVKAEBEAVAIANKSDFGLQSSVFTPRDOKAF 418
      397  MRPCVLTNCDDMTCKVEEIFFGVMSILSPDTEAVLERANDTTFGIAGVFTRDIOGRAH 456
Qy      419  DIANKLEVGTVHINNKTRGPDNFPFLGLKSGAGVQGIKRSIEAMTNVKSIVLDM 474
      457  RVAAELQAGTCFINN-YNVSVELPFGYKSGFGRENGRVTLEYYSQLTVCVEM 511

RESULT 51
US-09-949-016-7505
; Sequence 7505, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 7505
; LENGTH: 553
; TYPE: PR1
; ORGANISM: Human
; US-09-949-016-7505

Query Match      25.2%; Score 602, DB 4; Length 553;
Best Local Similarity 31.1%; Pred. No. 1.2e-47;
Matches 148; Conservative 99; Mismatches 197; Indels 32; Gaps 8;

Qy      23  LSPIDSSL--GFPAMTR-----EEVDHAKAGREALPAAALTVVERAQLTK 70
      79  VERPADSGTRKAEPEPTGRVIATFTGSGKEVNLAVQNAKAARKISQSGMERCKILB 138
Qy      71  AADIIRDRBEIATVLAKESIKAYNASVTEVVTADLIRYAAEEGIRLSTADEGSKMDA 130
      139  AARIIRERDEIATMGCINNKSIFEARLIDISWQCLEYYA-----GLAASMAEHIQ 192
Qy      131  STGHKLAVIRROPGVIVLAIPYNYVNTSGSKIAIPALLIGANVMFKPTQSGVSGVIA 190
      193  LPPGSGFYRRREPLGVCVIGANNYFPFOIASWKSAPALACGNAMVFKPSFTFVSALLIA 252
Qy      191  KAPFAEGLPAGVNTITGRGSEIGDYIVHEEYVNFINTGSPVGRIGTLA--GMRPLM 248
      253  EYISEGVPGLNNAVQ--GAATGQPLCGHPDAKVSFTGSPVTKIMEMSKGIKPT 311
Qy      249  LELGSGAGIVLADADLDNNAKOIVAGAYDSQRCITAKRVLVVEEVADELAEKISENV 308
      312  LELGSGSPLIIFSDCMNNAVKGALMANFLTOGQVCNGRVAVQKEIIDKFTBEVYVKQT 371
Qy      309  AKLSVDP--FDNATVTPVIDNSADTIESLVDAKQKAKEL-----NEFKRDGRU 358
      372  ORKIDPPLLEDTRMRPBLINRPHLEVLGVPVKAAGKAVLGGDIYVEDEPKLKDGY 431
Qy      359  LTFGLFDHVTLDNKLAMEEPFGILPIIRVKAEBEAVAIANKSDFGLQSSVFTPRDOKAF 418
      432  MRPCVLTNCDDMTCKVEEIFFGVMSILSPDTEAVLERANDTTFGIAGVFTRDIOGRAH 491
Qy      419  DIANKLEVGTVHINNKTRGPDNFPFLGLKSGAGVQGIKRSIEAMTNVKSIVLDM 474
      492  RVAAELQAGTCFINN-YNVSVELPFGYKSGFGRENGRVTLEYYSQLTVCVEM 546
```

```
RESULT 52
US-09-710-279-1030
; Sequence 1030, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348005
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1030
; LENGTH: 496
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1030

Query Match      25.1%; Score 601, DB 4; Length 496;
Best Local Similarity 32.5%; Pred. No. 1.2e-47;
Matches 152; Conservative 95; Mismatches 189; Indels 32; Gaps 10;

Qy      6  QNVNGEW--KSNVQIRILSPIDSSLGFPVAMTRVEVDHAKKGRALP--AMALTV 61
      10  RQYIDSEWSSNKTTRDILINPYNOETITVAEGTKEVERAILAARRFEGENSLETS 69
Qy      62  YERAOYLHKAADIIRDRKEIATVLAKESIKAYNASVTEVVTADLIRYAAEEGIRLSTS 121
      70  EYRGKKVRAVADKIKENRELAKEFTLDGTLESYADMDIHHVVFYA--GL--AD 124
Qy      122  ADEGSKMDASTGHKLAIVROPGVIVLAIPYNYVNTSGSKIAIPALLIGANVMFKPTQ 181
      125  KGGELINSPIINNAASKVKEPVGAVTQITPNVYPLQASWXTIAPALATGCSLVNKPSEI 184
Qy      182  GSVGLVLAKAPFAEAGLPAGVFNTITGRGSEIGDYIVHEEYVNFINTGSPVGRIGTL 241
      185  TPLTIRVLEMBEYGFPGKGTINLVGAGSEVDWMSGHEEDVLSFTGISTGKIMQ 244
Qy      242  AG--MRPIMLELGSGAGIVLADADLDNNAKOIVAGAYDSQRCITAKRVLVVEEVADE 299
      245  AANHVTVALRELGNQNPNIIPDDADPELAVDQALNGYFHAGQVCSAGSRILVHNDIKQ 304
Qy      300  LAEKISENVAKLSVQDPDNAT--VTPVIDDNSADTIESLVDAKQKAKELNEFKRDGRU 358
      305  FEKALIDRVSKIKLNGFPQDTEMGPVISTARDKIEGYMEVAKKOGATIALIGRRPERE 364
Qy      359  -LTFGLFDHVTL-----DMKLAMEEPFGILPIIRVKAEBEAVAIANKSDFGLQSSVFT 411
      365  DLQAGLFEPEYITTCDSMRIVQBEVSGPVTVGCFADDEEALRLANDSITGLGAIY 424
Qy      412  RDKAFADIANLKEVGTVHINNKTRGPDNFPFL-----GLKSGAG 453
      425  KQIGRAQVRANKLKTGTWIN-----DFHHYFAQAPWGVYKQSGIG 465

RESULT 53
US-09-134-001C-4383
; Sequence 4383, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn DoucetRe-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
```

```

? PRIOR FILING DATE: 1997-11-08
? PRIOR APPLICATION NUMBER: US 60/055,779
? PRIOR FILING DATE: 1997-08-14
? NUMBER OF SFO ID NOS: 5674
SFO ID NO 4383
? LENGTH: 506
? TYPE: PRT
? ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4383

```

Query Match	25.1%;	Score 601;	DB 3;	Length 506;
Best Local Similarity	32.5%;	Pred. No. 1.2e-47;		
Matches 152;	Conservative 95;	Mismatches 189;	Indels 32;	Gaps 10;

Qy	6	ÖNŸVNGEM--KSSVNÖIEILSPIDBSLGFVAMRREEDHAMKGRALP--AMAAITV	61
Db	20	RÖYIDGEMWESSKNTNRDIINPNÖGTITFTVABGRTKEDVERALLAARSFEDGMSLETS	79
Qy	62	YERÄQYIAHKAADIIEEDKEBIATVLAKESKAYMASVEVATFADLIRYAAEGIRLSTS	121
Db	80	EVGKKURAVADTIKENREBELALETLDGKLTLESYDMDMDINHVMFYFA--GI--AD	134
Qy	122	ADEGKKMDASTGKKLIVIRROPVGIYLAAPNYNPUNSGSLAPALIGAVNMFKEPTQ	180
Db	135	KDGEIINSGTIPAAESKYVKEPVGVTQITPNMYFELQASMKIAPALITGSLVWK8SEI	194
Qy	182	GSVSGVLAKAFAPAEAGIPAGVFNFTITGRGSEIGDYIYHEEYFNINFTGSTPVGORIGL	241
Db	195	TPLTTTRVFEMEEVFPFKGTINLVLGASSEVGDVMSGHEVDLVSFTGQIETGKIMKQ	255
Qy	242	AG--MRIMLEIGKXAGIYLAADADIDNNAKOIVAGADYDSQORTAIKRVLYVBEVADE	295
Db	255	AAHVITDVALEIGKKPNIIIFDADBELVADALNGGFPHAQOVOSASRLLIHNHDIKD	314
Qy	300	LAEKISENVAKLSVGPPDNAT--VTEVIDNSADFIESIYVDAROKAKEINFEKRODL	358
Db	315	PEKALIDRVSKITLGNFGPDOTDMGVITSTARDKIEGMEYAKKOGATTIAGKKRPRE	374
Qy	359	LTFEGLFEDHVTL-----DMKLAEBEPFGRILPIIRVDAEBAVAIANKSDFGLOSSVFT	411
Db	375	DLQAGFFPEPTVITDDCOTSMRIYQOEVEVFPVATVEGPADEBEAIRLANDSIYGLAGMFT	433
Qy	412	RDOQKAPDANKLEVEGTVIHNKTKGSGDPNPPFL-----GIKSGGAG	453
Db	435	KDJKKARVANKLGLTWIN-----DEHFYPAQAPWGSGYKOSGIG	475

RESULT 54
US-09-489-039A-13820
; Sequence 13820, Application US/09489039A

```

: APPLICANT: Gary Breton et. al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
: TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 2709.2004001
: CURRENT APPLICATION NUMBER: US/09/489,039A
: CURRENT FILING DATE: 2000-01-27
: PRIOR APPLICATION NUMBER: US 60/117,747
: PRIOR FILING DATE: 1999-01-29
: NUMBER OF SEQ ID NOS: 14342
: SEQ ID NO 13820
: LENGTH: 514
: TYPE: prt
: ORGANISM: klebsiella pneumoniae
: US-09-489-039A-13820

```

```

Query Match      25.1%; Score 600.5; DB 4; length 514;
Best Local Similarity 31.0%; Pred. No. 1.4e-47;
Matches 148; Conservative 89; Pmismatches 226; Indels 15; Gaps 6;

QY      4 EYQNVNNEKSSVNVQ--LEILSPDDSLGCVNPATREYVHAKKAGREALPAAALTY 61
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

Dh 4. QHBMVYIDGQVSGRGGMQIDVLPNTPATEALLSRIPDGTAKEARLMDAABRQPCWEMALPA 100
Qy 62 YERAOYLHRAADIIEERDKEIATVLAKEISKANAASTEVYRTDILIRYAAESIRLSTS 121
Db 101 IERAGWLRRIAGIRORADEIAGLIVAEGGKIQOLAAVEVAFTADYLDYMAEWARFY--- 157
Qy 122 ADRGGMMDASTGHKLAVIRORPVGIVLMAIPANYPVNLSGSKIPALIGGVNWFKPTQ 181
Db 158 --BGEIYOSDRPENILVFKALGOTTGLLEWNPFFLIRAKLAPALITGTITYIKSER 215
Qy 182 GSVSGVLAKAPABAGLIPAGVENTITTRGSEIDGYIEHEEVNFINTGSTPVGORIGKL 241
Db 216 TPNNALAFABEIYHQVGLPKGVFNVLTRGGETVQGEGLANPKVAMVSMTGSVAABEKITMAA 275
Qy 242 A--GMRPIMLELGKQAGIVLADADLDMAAQIYAGADVYSGORCTAIKRLVYEEBVADE 299
Db 276 AAKNIITKVCIELEGKAPALVMDADDELAIVAVVDSRVINSQVCNCEERYVVOOGGIYDR 335
Qy 300 LAEKISENVAKLSVGPDP--DNAUTPVYIDNSADPISLVADAROKAKEL---NEFKR 354
Db 336 FVNRLEGMAKAVQFGPBPATRDIMAGPLIMAAARQVAGKAKAVAGCARVALGOGPLEG 355
Qy 355 DGRLLTPGLFDHVTLLDMKLAMEEPGPIILPIIRVKDAEAVAIAKNSDPFGLQSSVFTTRDF 414
Db 396 KGFYFPTLLIDVROEMDIIEETGFGPVLPVAFESTLDEALAMANDSDYGLTSSITYTRDL 455
Qy 415 QKAFDLANKLEVGTVINNKTRGSDNPPPLGLKSSGAGVQGIIRYSIEAMTNVYSIYL 472
Db 456 NVAMKAIKGGIKETETIYNRENFEAMQGF--HAGMRSKSGIGGADGHHGIANEYIQTQVTVL 512

```

RESULT 55
US-09-538-092-1147
; Sequence 1147, Application US/09538092

```

: GENERAL INFORMATION:
: APPLICANT: Glact, Loic
: APPLICANT: Mansfield, Traci A.
: TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
: FILE REFERENCE: 15966-542
: CURRENT APPLICATION NUMBER: US/09/538,092
: CURRENT FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: 60/127,352
: PRIOR FILING DATE: 1999-04-01
: PRIOR APPLICATION NUMBER: 60/178,965
: PRIOR FILING DATE: 2000-02-01
: NUMBER OF SEQ ID NOS: 1387
: SOFTWARE: CurataseqFormatter Version 0.9
: SEQ ID NO 1147
: LENGTH: 493
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (0)..(0)
: OTHER INFORMATION: Polypeptide Accession Number P49189
: US-09-538-092-1147

```

Query Match	25.0%;	Score 599;	DB 4;	Length 493;
Best Local Similarity	31.2%;	Pred. No. 1.8e-47;		
Matches 145;	Conservative 98;	Mismatches 191;	Indels 30;	Gaps 7

OY	33	FVPMR-----EVDHAKAGSEALPAMALIVYRAOYLHAAIIIRDREI	82
Db	31	FEPATRVATPTCGSEKEVLNAYQNAKAPKMGQSGMERCHILLEANIIREREI	90
OY	83	ATLAEISAKYMASTEYVRADLDLRVAEGIRLSTSDGEGMDASTCHKAVLRQ	142
Db	91	ATMGCSNNGKSTFEALRDLDISWCSEYTA-----GLAASMAEGHILQPGSGYTRRE	144
OY	143	PVGIVLAIAPYNPVNLGGSKIAPALLIGQNVMEFKPTQGSVSGVLVAKAPAEAGLPAGV	202
Db	145	PLGACVQIGAMNVPFQIATWKSAPPLACGAMNAVFDPSPFTFVSALLLAIYISEGVPGL	204


```

; TITLE OF INVENTION: 32140, A No. 6569657el Human Aldehyde
; FILE OF INVENTION: Dehydrogenase
; FILE REFERENCE: 35800/205243
; CURRENT APPLICATION NUMBER: US/09/717,926
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-717-926-2

```

```

Query Match      25.0%; Score 597.5; DB 4; Length 923;
Best Local Similarity 30.8%; Pred. No. 6.9e-47;
Matches 152; Conservative 97; Mismatches 207; Indels 37; Gaps 13;

```

```

QY 5 YONYVNGEMKSSVN--QIEILSPIDDSLGFPVPMTRREVDHAMKAGREALP--AWAALT 60
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 443 YCCFINGCFTDADGCKTVDPTINPTDSTICKVSYASLADVDKVAALAKAFENGEMGRN 502
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 61 YVERAQYLKADIIERDKEEIAATVLAKEISKAYNASV--TEVVRADLIRYAAE--EGIR 117
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 503 ABERGLMAYRLADILEENQEBELATIEALDSGAVYTLAKTHIGMSVQTRFYFAGWCDKIQ 562
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 118 LSTSADEGKMDASTGHKLAVIRROPVGIYLAIPNYPVNLGSKIAIPALLIGSNVMMK 177
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 563 GSTIPIIN---QAPRNRLFTTKEPLGVCAIIPNNYPMMLAMSAACLAAGNTLVLK 618
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 178 PPTQGSVSGVLVAKAPAEAGLPAGVFNITIGRSGEIGDYIVHEEVNFINTGSTPVGOR 237
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 619 PAQVPLTLTKRELSEVKAQFPKGVNIIIPGSGIAGQRLSEHPDIRKLGFTGSTITIGQ 678
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 238 IGR--LAGRPIMLELGGKAGIVLADLDLNAKQIVAGAVDYSQGRCTAIKRVLYVE 294
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 679 IMKSCAVSNLKKVSLGKSPILIPNDCELDKAVMGWGAVPFNKGNCIAAGRLFEVE 738
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 295 EVADELAEKISENVALLSGDPPDNATVTPVD--DNGADFIESLVNDAQKAKE--- 348
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 739 SIHDEFVTYVVEIKMKIKGIDPLDRST---DHGPNKHALEKILQYCEFGVEGAT 792
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 349 ---LNEFRKDGSLILPGFDHYTLDMKLAMEEPFPIILIRVK--DAEEAVAILANKSD 402
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 793 LVYCGQVQRPGEFMEPTVFTDVEDYMLAKESFPPIVVISKFGQGDIDVILGRANSTE 852
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 403 FGLQSSVFTFRDFOKAFDIANKLEVGTVHIN--NKTGRGPDNPFPLGLKSGAGVQIRYS 460
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 853 YGLASGVFTRDINKAMVYSEKLEAGIVFINTYKTDVAA---PFGVKGSGFGKDLGSEA 909
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 461 IEAMNVKSIVD 473
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 910 LNEYLTKTIVTLE 922
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||

```

```

RESULT 59
US-09-902-540-15889
; Sequence 15889, Application US/0902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15889

```

```

; LENGTH: 477
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-15889

```

```

Query Match      24.9%; Score 595.5; DB 4; Length 477;
Best Local Similarity 32.0%; Pred. No. 3.7e-47;
Matches 154; Conservative 86; Mismatches 224; Indels 17; Gaps 6;

```

```

QY 1 LTKYQYVNGEMKSSVNOIILSPIDDSLGFPVPMTRREVDHAMKAGREALPMAALT 60
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 2 LAERYPYLANRPQPNALAVTHKYSGEVTHAVADAAVEEILAAVBAEPMKLA 61
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 61 YVERAQYLKADIIERDKEEIAATVLAKEISKAYNASVTEVVRADLIRYAAEGRILST 120
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 62 PYARQAVLEHCVRREBAEELALALCTEAGKPLRDAGVEDRLDITKAAAEAVR--- 118
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 121 SADEGKN-----DASTGHKLAVIRROPVGIYLAIPNYPVNLGSKIAIPALLIGSNV 174
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 119 ---GGGEVLNLEVSRAAGYR--GFTQVVPVGPSPFTFPNPLNVAKHVAPALAAQCP 174
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 175 MEKPTQGSVSGVLVAKAPAEAGLPAGVFNITIGRSGEIGDYIVHEEVNFINTGSTPVP 234
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 175 ILKPSDRTPPVSALEIAETELPEGAFSVLPVRLIEDIGP-LIEDRLKLSFTGSEKV 233
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 235 GQRIKLAGMPIMLELGGKAGIV--LADADLNAKQIVAGAVDYSQGRCTAIKRVLYV 292
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 234 GHELKARARGRKKVYVLELGSNAACVVDENRGAFLDFIDRVAQAGAFQAGQSCISQRLV 293
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 293 VEEVADELAEKISENVALLSGDPPDNA--TVTPIVDNSADFIESLVNDAQKAKEINE 351
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 294 HASLYDALREKLIVERAKALRPGHPDEATTLGPMIDEPBARRLBEGIOAAVQGRARVL 353
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 352 FRDRLRLTPGLFDHYTLDMKLAMEEPFPIILIRYVDAEENVAIALKSPGLQSSVFT 411
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 354 GGRRSVLDATVYLEGVBPDAISSEARGPVLLQPFDSFBEALAQVNSGRFGLOAGLFT 413
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 412 RPFQKAFIANKLEVGTVHINNKTRGPDNPFPLGLKSGAGVQIRYSIAMTVKSI 471
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 414 QDLSAAMKAMBELEVGIVGVGVDFSRVDYMPYGVKSGSIGRECVKAIADMTLRLV 473
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 472 L 472
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 474 L 474
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||

```

```

RESULT 60
US-09-328-352-7337
; Sequence 7337, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7337
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-7337

```

```

Query Match      24.8%; Score 594; DB 4; Length 505;
Best Local Similarity 31.1%; Pred. No. 5.6e-47;
Matches 153; Conservative 109; Mismatches 198; Indels 32; Gaps 11;

```

```

QY 1 LTKYQYVNGEMKSS--VNOIILSPIDDSLGFPVPMTRREVDHAMKAGREALP--AWA 57
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 24 MARQGHPIIDKLVAAELDKVDIVNPSTEVVAQISIGSODVSAVSAEHAQONAMA 83
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 58 ALTVERAQYLKADIIERDKEEIAATVLAKEISKAYNAS--VTEVVRADLIRYAAEG 116
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||

```

```

Db      84  ETTPEYRGVLTAKLADLIBOYGEELAQLETLSTGKLINISRHLEVAQSVFLVFPAGMAT 143
      117  RL-----STADEGKMDASTGHKLAVIRROPVGIATLADPYNYPVNLSSGKIAPALI 169
      144  KINGOTMOPISIPMOCEKTYAFT-----LRQPVGVVAGIVPMNFSIMIGWKISALT 196
Qy      170  GGVVWVFKPPTQSSVGLVLAKAFAEAGLPAGVFNITITGSGSEIGDYIVHEEVNFINT 229
      197  TGGTIVLKPEEPASLSLRLAEALIEAGIPAGVINVTGKG-DTGQVLIESPLVKVSFT 255
Qy      230  GSTPVGQRIKTLA---GMRPIMELGKDGAGIYLADADLNAKAQIVAGAYDVSQGRCTA 286
      256  GSVPTGAIAGKLMASDILTVSLELGKNAIAYLADANDEILPTLIQATFPHQGYCAS 315
Qy      287  IKKVLVEEVDLAEKISENVAKLSVDPDNATV-TPVIDNSADFIESLVVDAROK- 344
      316  PERFFVHRTKYDELVKLSKALSQFKISAMDEGSMGFLSNQPHFKVGHYLDMAKANN 375
Db      345  ----GAKELNEFRDGLLTPLGFDHVTLDMKLAMEBPFGPIPIIRVKDAEAVAIANK 400
      376  QIIAGEEALDQI---GTFVQPTLISFRKNTDPLFSEBTFGPVVGVPFETDEELIQLMNQ 432
Qy      401  SDFGLQSVPTRDFOKAFDIAKLEVTGVIHNNKTKGRGPNPFLGKSGAGVQGIARS 460
      433  SRRLGASTITNDLSKALRLIPKIEAGTLVNNHTFLDP-SVPEGVKASGIGREPSDAF 491
Db      461  IEAMTNVKSIVL 472
Qy      492  IEDYTELKSVMI 503

```

RESULT 61
US-09-252-991A-28449
; Sequence 28449, Application US/09252991A
; Patent No. 6551795

```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28449
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28449

```

```

Query Match      24.7%; Score 591.5; DB 4; Length 748;
Best Local Similarity 30.3%; Pred. No. 1.8e-46;
Matches 148; Conservative 91; Mismatches 215; Indels 35; Gaps 7;

Qy      6  ONVYVNGKSSVNOIEILSPIDSSLGFPVPMTRREVDHAMKAGREALPAMALITYEPA 65
      265  KMINIREVESKQVFNYPNATSELIGEVASGGAELDAVAARAEFPMAATPAKERA 324
Db      66  QYLAHKAADIIERDKEIATVLAKESKAYNASYTEVV-RTADLIRYAABEGRILSTADE 124
      325  RLHRRIRGELIDRVNPHLALETDTGLPIHOTKGVILPRASHNFEFFAIVCTRMN----- 379
Qy      125  GCGMDASTGHKLAVIR-----QPVGIYLAAPNYPVNLSSGKIAPLIGCNVVMK 177
      380  -----GHSYVDQOMLNTYLYQPVGCVLSPMNVNVPMTATWKTAPLALGNRAVAK 431
Db      178  PPTQSSVGLVLAKAFAEAGLPAGVFNITITGSGSEIGDYIVHEEVNFINTGSTPVGOR 237
      432  MSLSLPLTANELGRVLAHEAGIIPGVIVNVVQGYASAGDALVIRRDVRAVSFTGSTATGR 491

```

```

Qy      238  IGTLAGMRPIMELGKDGAGIYLADADLNAKAQIVAGAYDVSQGRCTAIKRVLYVEVA 297
      492  IEMAGIKIKYSMELGKSPVLVEEDALRALDAAFTFLSINGERCTAGSRIFVQESYV 551
Qy      298  DELAEKISENVAKLSVGDPEFNAT-----VTPVIDNSADFI-----SLVVDAROK 345
      552  PQEVAEFAPARARLIVGDQDPRTQVGSMTTOAHDKVTGYIRIGLEBAGITLVAGLEPR 611
Db      346  AKELNEFKDGRLLTPGLFDHVTLDMKLAMEBPFGPIPIIRVKDAEAVAIANKSDPGL 405
      612  A-CLPMLSKGQPIQPTVADVNRRIRIAQEBELFGPVVCLIPFKDBAALRLANDVEYGL 670
Qy      406  QSSVPTRQKAFDIANKLEVTGVIHNNKTKGRGPNPFLGKSGAGVQGIIRYSIEAMT 465
      671  ASYIWTQDIGKHLRLARGIEAGVVFINSQVNDLRO-PFGGVYASGTGREGGEYSFEVPA 729
Db      466  NVKSIYLDIM 474
      730  EIRNVCTISM 738

```

RESULT 62
US-09-328-352-6757
; Sequence 6757, Application US/09328352
; Patent No. 6562958

```

; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6757
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6757

```

```

Query Match      24.7%; Score 590.5; DB 4; Length 500;
Best Local Similarity 33.2%; Pred. No. 1.2e-46;
Matches 161; Conservative 76; Mismatches 225; Indels 23; Gaps 10;

Qy      1  LTKYQYVNGKSSVNOIEILSPIDSSLGFPVPMTRREVDHAMKAGREAL--PAMAA 58
      12  KTDADFLVYINGQFESGAATFESINPATGCVMAQMPPEARTEDEVNRAVQAASQLKASAWAG 71
Db      59  LTVYERAOYLHKAADIIERDKEIATVLAKESKAYNASYTEVVRTADLIRYAABEGRIL 118
      72  LTASQRGKLYKLADLIBERKAPQLAQIERTSDGKIIIRETSQIAVVAEYRYTA--GI-- 127
Qy      119  STSAD--EGKMDASTGHKLAVIRROPVGIYLAIPYNYPVNLSSGKIAPLIGCNVVMF 176
      128  ---ADLRESPLIPDKADMQAMIVBPPGVVAIVAPMNSQLLSAVKVPALAAAGTIVL 184
Qy      177  KPPTQSSVGLVLAKAFAEAGLPAGVFNITITGSGSEIGDYIVHEEVNFINTGSTPVQ 236
      185  KASBQPGGLLAFALVHAGPAGVAVNITFGEGCAVLSHDPVHVAATGTGGETAR 244
Db      237  RI--GKLAGMRPIMELGKDGAGIYLADADLNAKAQIVAGAYDVSQGRCTAIKRVLYVE 294
      245  HTVNSAENILAVSLELGKSPFIYFADADIQSAVNAQVAALFAATGSCVAGSRLVBE 304
Qy      295  EVADELAEKISENVAKLSVGDPEFNAT--VTPVIDNSADFIESLVVDAROKAKEL---N 350
      305  SIKDEFQVLVRVOTIKGLPHEMATERGPCTIAROKOKIQOVVASSIQGAKLLTGSK 364
Db      351  EFKDGRLLTPGLFDHVTLDMKLA--WEPPGPIPIIRVKDAEAVAIANKSDPGLQS 407
      365  SLERAGYVYPITLD--CTDVPQADCIITELFGPVLSDVSFQDBAEOVKANSTPYGLAA 422
Qy      408  SVPTRDFQKAFDIANKLEVTGVIHNNKTKGRGPNPFLGKSGAGVQGIIRYSIEAMTV 467

```

Db 423 GFFTNLSAHMRTKVGSGIWLNTYRAVSP-LAPFGHGSLGHGREGANAVLDYTTT 481
QY 468 KSIVL 472
Db 482 KTWVL 486

RESULT 63

US-09-543-681A-4515
Sequence 4515, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:

APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 4515
LENGTH: 499
TYPE: PRT
ORGANISM: *Proteus mirabilis*
US-09-543-681A-4515

Query Match 24.4%, Score 585, DB 4, Length 499,
Best Local Similarity 28.4%, Pred. No. 3,9e-46,
Matches 139, Conservative 109, Mismatches 204, Indels 38, Gaps 9,

QY 8 YVNGEM---KSSVNOIEILSPIDSSLGFPAMTTEEDVHMKAGREALPMAALTVE 63
Db 18 YIHGGVDDAQPECEBPFAINPATGEVIAHLOSAGEDINWAVESAKOGOKTMAAMPAA 77
QY 64 RAOYHKAADIIRDKKEIATYLAKEISRAYNAS-VTEVTRADLIRYAEEGIRLSTSA 122
Db 78 RSRILRRADILIRERDELALHLETTDTPKPLSETRVYDIWVGADVLEYAA----- 127
QY 123 DEGGKMDASTGHKLV-----IRQPGIVLAIAPYVNVLSGSKTAPALIGNVYM 175
Db 128 ---GLIPALEGQIPIRDSAFVYTRREPGLVAGVAGAMVPIQIALMKAPALAAGNAMI 184
QY 176 FRPPTGVSGLVLAFAAEAGLPAGVNTITGRGSEIGDYIHEHEVNFINTGSTPGV 235
Db 185 FRPSEMTSLTKLTAETIYBAGLPQGVFNVVG-DATVQWLTQHDIAKVSFTGSDTG 243
QY 236 QRI---GKLAKMRPIMLELGGKDAIVLADADLDNNAKOIVAGAYDYSQORCTAIRVLV 292
Db 244 KKVMASSASSSLKDVMTLGGKSPILIIDDADDDTAADIAMMANFYSSGQVCTNGTRFV 303
QY 293 VEEVADLEAKISENVAKLSVGPDPG-NATVTVVIDNSADPESLVNDARQKAELNE 351
Db 304 PEKLKQAFBAKIIERVSRIKRVGSPLEMDNFGPLVSPHLEKVLSTYELGKAGALTLG 363
QY 352 PKR-----DGRLLTPGLFDHYTLPMKLAEEPPPIIIRVKDAEVAALANKSDRG 404
Db 364 GERLPAGALAKAYVAPYPTFTCTDMDKITDDEIFGPVMSISLYQDEYIIRANNSVYG 423
QY 405 LOSVFTTRDFOKAFDIANKLEVGVTHINNKTRGPNPFLGLKSGAGVQGIIRYSIEM 464
Db 424 LAAGLVTRDLTYHRYIHOLEAGICWINT-WGESAEMPVGYKSHGSGVGENGLVTLQNY 482
QY 465 TWKSTVLDLM 474
Db 483 TQIKSTQVEL 492

RESULT 64
US-09-134-001C-4246
Sequence 4246, Application US/09134001C
Patent No. 6380370

GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4246
LENGTH: 488
TYPE: PRT
ORGANISM: *Staphylococcus epidermidis*
US-09-134-001C-4246

Query Match 24.4%, Score 583.5, DB 3, Length 488,
Best Local Similarity 29.9%, Pred. No. 5.1e-46,
Matches 147, Conservative 95, Mismatches 214, Indels 35, Gaps 11,

QY 1 LTKY-----QNVNGEMKSSVN-QIEILSPIDSSLGFPAMTTEEDVHMKAGRE 51
Db 7 LRSEYKRNKNTKQYINGEWDSAGETIDVINPTEVWKIAGNEEDVNAKADADK 66
QY 52 ALPMAALTVEERAQYLAHKAADIIRDKKEIATYLAKEISRAYNASVTEVTRADLIRYA 111
Db 67 VLEFRSHSVBERRLDLKTYKEYNRKNDDLEAITDELGAPLVS-----ENVHY- 117
QY 112 AEEGIRLSTAD---EGGKMDASTGHKLVIRROQVIGVLAIAFYVNVLSGSKTAPAL 168
Db 118 -QMGINHFTTAARDALDSFQFEORGDDLVV--KEAIVAGVLTVPWNPTNQTSLKLAAP 174
QY 169 ICGNVNMPKPTQGSVSLVLAFAAEAGLPAGVNTITGRGSEIGDYIHEHEVNFNF 228
Db 175 AAGSPVVLKPESEETPFAAILTAEIFDKVGPVGFNLVNGSGSGGNPLSEHPKRMMSF 234
QY 229 TGSSTVGORIGLAG--RPIWLELGGKDAIVLADADLDNNAKOIVAGAYDYSQORCTA 286
Db 235 TOSGPTGSKIMEKAKDKRYKSLBLGKSPYIVLDDVVEBAANAATTKKVANNQCVCTA 294
QY 287 IKRVLVEEVADELAEKISENVAKLSVGPDPDNAT-VTVVIDNSADPESLVNDARQK 345
Db 295 GTRVLIPRSIKEDYLTVAKEAFSKVKGQPREBGTQVPIISKQPDQVQYIDKGINEG 354
QY 346 AKEL-----NEPKDGLLTPGLFDHYTLPMKLAEEPPPIIIRVKDAEVAALAN 399
Db 355 AELFYGGGKEPEGLDKGYFARPTTFINVDNMHTIAQBEIFGPVMSITYNNLDEAIBIAN 414
QY 400 KSDPGLQSSVFTTRDFOKAFDIANKLEVGVTHINNKTRGPNPFLGLKSGAGVQGIIRY 459
Db 415 DTKYGLAGVIGKDDTIRHARVSIKIEAGTIEL-NEAGKRPD-LPFGYKESGLGEMKDY 472
QY 460 STEAMTVKSI 470
Db 473 GIEBFLVYSI 483

RESULT 65
US-09-489-039A-9023
Sequence 9023, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342

QY 238 DELAEKISENVAKLSVGDPEF-DNATVTPVIDNDSADFIESLVVDARQKAKELN---EFK 353
DB 301 EALISOKLADINAQAQVGDMASSSTTEGFPQONQYVVKALIAEALQAKALKSCQKLP 360
QY 354 RRGRLTTPGLFDHVTLDMLKAMEEPPGPIIPITIRVDAEVAALANKSPGLOSVFTRD 413
DB 361 EGGYIAPILITREVSOSCSRVOEBOGPVLPVAKTIDINDALAPANDSEFGGSIWSSD 420
QY 414 FOKAFDIAKLEVEGVTHINNKTGRGPDNFPPLGLKSGAGVQIRYSIEM 464
DB 421 LKTAQTYATQLOCGTYWITHTAEVLP-HAPFGWMKMSGLGAE---FGLEGL 467
RESULT 68
US-09-344-882-20
; Sequence 20, Application US/09344882
; Patent No. 6764851
; GENERAL INFORMATION:
; APPLICANT: Nikolau, Basil J
; APPLICANT: Wurtzle, Eve S
; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; APPLICANT: Schnable, Patrick S
; APPLICANT: Ke, Jinsan
; APPLICANT: Johnson, Jerry L
; APPLICANT: Allred, Carolyn C
; APPLICANT: Faltland, Beth
; APPLICANT: Lutziger, Isabelle
; APPLICANT: Wen, Tsui-Jung
; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
; TITLE OF INVENTION: Acetyl CoA Levels in Plants
; FILE REFERENCE: 201573
; CURRENT APPLICATION NUMBER: US/09/344,882
; CURRENT FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,717
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 20
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
US-09-344-882-20
Query Match 23.8%; Score 568.5; DB 4; Length 538;
Best Local Similarity 32.0%; Pred. No. 1.5e-44;
Matches 156; Conservative 85; Mismatches 208; Indels 39; Gaps 13;
QY 6 QNYVNGEMKSVN--QIEILSPIDSSLGFPVPMTRREVDHAMKAGREALP--AMAAALT 61
DB 59 QLLINGNFVDSASGTFPTLDPRTGEVIAHVAEGDAEDINRAVKAARAFDEGPPMKSA 118
QY 62 VERAOYLKAAADIIRDKKEIATVLAKEISKAYNASVT-EVVRTADLIYYA--AEGR 117
DB 119 YERSVTLRFADLVEHSEELASLETWNGKPYOQSLTAIEPMFALFFYVAGMDKIHG 178
QY 118 ILSADEGKMDASTGHKLAVTRROPVGIIVLAIPYNPVNLSGSKIAPALIGANVMWK 177
DB 179 LTIPLAD--GNYOVHTLH-----BPIGAGQIIPNPFILPMFAMKVGALACGNTIYVK 229
QY 178 PPTQSVSGVLAKAPAEAGLPAGVFNITITGRSGEIGDIYVHEEVPNFTGSPVQOR 237
DB 230 TAEQPLPFTFVAGKLFLEAGLPFGVNLIVSGFATGALASHMDVDKLAFTGSDTDGKV 289
QY 238 IGLKA---GMRPIMLELGKAGIYVLADDLNNAKQIVAGAYDVSQGRCTAIKRYLVE 294
DB 290 ILGLANSMLKEVTLLELGKSPFIYFEDADIDKAVELAHFALFNGQCCGAGSTRFVHE 349
QY 295 EVADELAEKISENVAKLSVGDPEFDNATVT-PVID-----DNSADFIESTLVVD--ARQKAK 347
DB 350 KYIDFVEFKSKRALKRVVGDPEFRKIGEGPOIDLKQFEKWMKTIYISGIESNAITTECGSD 409
QY 348 ELNEFRDGRLLTPGLFDHVTLDMLKAMEEPPGPIIPITIRVDAEVAALANKSPDGLQS 407

DB 410 QIGD---KGYPLOPFVFSNVDDMLIADOLFEGVQSIKSPSDVEVKKRNETYVGIAA 466
QY 408 SVFTRDPQKAFDIAKLEVEGVTHINNKTGRGPDNF---PFLGKSGAGVQIRYSIEA 463
DB 467 GVFINLDTNARKVSALKAGTWNCF-----DVFDAAIPFGGYKMSGNGREKGIYSLNN 521
QY 464 MTNVSIV 471
DB 522 YLQIKAVV 529
RESULT 69
US-09-716-865-6
; Sequence 6, Application US/09716865
; Patent No. 6548292
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Micheal G
; APPLICANT: Nagarajan, Vasantha
; APPLICANT: Chen, Mario W.
; TITLE OF INVENTION: Bacterial Plasmid Having Genes Encoding Enzymes for the
; TITLE OF INVENTION: Degradation of Aromatic Compounds
; FILE REFERENCE: BCJ016 US NA
; CURRENT APPLICATION NUMBER: US/09/716,865
; CURRENT FILING DATE: 2000-11-26
; PRIOR APPLICATION NUMBER: 60/167,062
; PRIOR FILING DATE: 1999-11-23
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Pseudomonas CT14
US-09-716-865-6
Query Match 23.5%; Score 562.5; DB 4; Length 486;
Best Local Similarity 30.3%; Pred. No. 4.8e-44;
Matches 149; Conservative 92; Mismatches 220; Indels 31; Gaps 10;
QY 3 KEYONYVNGEMKSVN--QIEILSPIDSSLGFPVPMTRREVDHAMKAGREALP-WAAL 59
DB 2 KDIRNFINGEYVTNNSGTYEKRNPDVDSLIGVHEAQCPVEDAVAAARALANGPWGKL 61
QY 60 TVYERAOYLH-KAADIIRDKKEIATVLAKEISKAYNASVTEVVRTADLIYYAAEGR 118
DB 62 SVYRCANLDDGVVAEINRRPDPLOAEIADTGKRAHLASHDIDIRGAANFPIPTDTIONV 121
QY 119 STSADGKMDASTGHKLAVTRROPVGIIVLAIPYNPVNLSGSKIAPALIGANVMFKP 178
DB 122 STSEFEMKTPGKTARSYGV--RTPRGVIAIICPMNLPLIMTWKCGPAMACGNTVVVKP 179
QY 179 PPTQSVSGVLAKAPAEAGLPAGVFNITITGRG-SEIGDIYVHEEVPNFTGSPVQOR 237
DB 180 SEATPTSTYTLIGEVNKKVGPVPGVYVNVNGVNSAGSFLTAHOGVNSITTTGETKTGTA 239
QY 238 IGLKA---GMRPIMLELGKAGIYVLADDLNNAKQIVAGAYDVSQGRCTAIKRYLVEE 295
DB 240 IMKAGADIRVVSLELGKMAAVVPADCFENALATVIRSAFENGQVCLGTERIYVERP 299
QY 296 VADLEAEKISENVAKLSVGDPEFDNAT-VTPVIDNDSADFIESLVVDARQKAK----- 347
DB 300 IFDKFVSLKERRAAIKRGPFPDADTKIGPLVSKIHQKVLISYAKAKABEGANIVLAGGV 359
QY 348 -ELNEFRDGRLLTPGLFDHVTLDMLKAMEEPPGPIIPITIRVDAEVAALANKSPDGLQ 406
DB 360 PNMPPDLDGCGWVEPTITGTPLESSPIVREIIFGCGCHIQPDTTEEELVANNVNSPYLA 419
QY 407 SSVFTRDPQKAFDIAKLEVEGVTHINN-----KTGRGPDNFPPLGLKSGAGVQIRYS 460
DB 420 TSIHTQDISRSARLATOIEVGLCWINSWFLADLRT-----PFGGSKQSGIGREBGHLS 472
QY 461 IEAMTNVSIVL 472

Db 473 LEFYTELRNMI 484

RESULT 70

US-09-252-991A-32298
 ; Sequence 32298, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 32298
 ; LENGTH: 475
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-32298

Query Match 23.5%; Score 562; DB 4; Length 475;
 Best Local Similarity 33.3%; Pred. No. 5.1e-44;
 Matches 157; Conservative 81; Mismatches 195; Indels 38; Gaps 12;

21 EILSPIDSSLGFPVPMATREEVPHAMKAGREALP--AMALTYVERAOYLHKADIIRD 78
 15 DAIDPSNRLLAAVAACDAADVDAVAARAPDEGMARLAYERKRVLLRLAEILAH 74
 79 KEERATV---LAKESKAYNASVTEVTRADLRVAAEGRILSTADEGKMDASTGH 134
 75 REELALLDLSLMGKRPWDAWN---IDVPGAHVFAWAE-----SLDLTYQVAPTAQ 124
 135 K-LAVIRROPVGIYLAIPNYVNLGSGKIAPALIGANVVFKEPTQSGVGLVAKAF 193
 125 QTAATITRVPLGIYGVAVPNNPFLDMAAMKLALAGNSGVLPKAEQSPFSALRLAEIL 184
 194 AEAAGLAGVFNITIGRSEIGDYIVHEEVNFINFTSGTVPQ---RIGKLAMRPTIME 250
 185 LEEAGVEGVANVPGISEQAGKALGHPEDALVFSTGVGYSAQSLKOWLE 244
 251 LGKDGIVLADA-DLDNAKQIVAGAYDYGORTAIREKLVVEEVADELAEKISENVA 309
 245 CGKSPVLVPAQDQDLDLALEKAFGLFFNQGSEVCSANSLVVERSHDFVRLAKAR 304
 310 KLSVGPEDNAT-VTPVIDNSADFIESLVVDARQKAKELNEFKDGRLLT----- 360
 305 DMQGPDLDPASBAGAVIDRRQIAGLIAIERAQGEATLLG---GGRQLTNGSDNFI 360
 361 -PGLFPHVTLDMKLAMEEPGPILPIIRVDAEBAVAIAKSPFGLOSSYFTDFQKAP 419
 361 EPTLFGVDRDMDLAREIFGPVALIASPDESEBAIRLANDSRYGLAASLMSDLRRAH 420
 420 IANKLEVTGVIHNNKTGRPD-NFPFLGKSGAGVQIGIRYIEAMTNVVS 469
 421 VARRIAGTVSV--TYDADVAVPRFGKQSGFGRDLSLHSFDKTYQTKT 469

RESULT 71

US-09-252-991A-25398
 ; Sequence 25398, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 25398
 ; LENGTH: 642
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-25398

Query Match 23.4%; Score 561; DB 4; Length 642;
 Best Local Similarity 31.7%; Pred. No. 1e-43;
 Matches 150; Conservative 79; Mismatches 186; Indels 58; Gaps 12;

31 LGFVPMTRE---EVDHAMKAGREALPMAALTYVERAOYLHKADIIRDKEIATVYL 86
 166 LGSTPLDADALMALDAVRAVDHQQGLWPNLRVARIQYVERFLARREQEAVVKLL 225
 87 AKESKAYNASVTEVTRADL-----RYAAEGRILSTADEGKMDAS 131
 226 MMEIGKLDKSDKEPDTCDYIVDTLEAKELDRSSRFLEQG----- 269
 132 TGHKLAVIRROPVGIYLAIPNYVNLGSGKIAPALIGANVVFKEPTQSGVGLVAK 191
 270 ---TLQIRRVPLGVALCKGPYNYPLNETFTTLIPALIGNTVVFKAQFVILLRPLLE 326
 192 AEAAGLAGVFNITIGRSEIGDYIVHEEVNFINFTSGTVPQRIKLAGMRP---I 247
 327 APRDS-FPGVGVINVIIGRRETVSALMASGKVDVFAFCTHSGASDLKGL-HRRPHRLA 384
 248 MELGKDGIVLADADLDNAKQIVAGAYDYGORTAIREKLVVEEVADELAEKISEN 307
 385 ALGLDKNGIYLPQVDLNAVNAEAVTGALSTFNGRCTLKILFVHESYLRPFLDKFSAR 444
 308 VAKLSGDPED-NATVPIIDNSADFIESLVVDARQKAKELNEFKDGR--LTPGLF 364
 445 LSLKPGMEWGVALTPLEPGKVDYLEGGLNDALAKARVNNAGGEGHRSFFYPALL 504
 365 DHVTLDMKLAMEEPGPILPIIRVDAEBAVAIAKSPFGLOSSYFTDFQKAF 418
 505 SPVSPMRIRYHEQGRPLPVPRYRELDVIDYVLRSDYQGLSLFGNDPAQIGRLVDAF 564
 419 DIANKLEVTGVIHNNKTGRPDNFPFLGKSGAGVQIGIRYIEAMTNVKSIV 471
 565 --AN-QVRIRIVNAQCQRGPDSPYPMGRKNSABGTMSVHDALRVF-STRTIV 612

RESULT 72

US-09-344-882-24
 ; Sequence 24, Application US/09344882
 ; Patent No. 6764851
 ; GENERAL INFORMATION:
 ; APPLICANT: Nikolau, Basil J
 ; APPLICANT: Wurtele, Eve S
 ; APPLICANT: Oliver, David J
 ; APPLICANT: Behal, Robert
 ; APPLICANT: Schnable, Patrick S
 ; APPLICANT: Ke, Jinsan
 ; APPLICANT: Johnson, Jerry L
 ; APPLICANT: Allred, Carolyn C
 ; APPLICANT: Fatland, Beth
 ; APPLICANT: Lutziger, Isabelle
 ; APPLICANT: Wen, Tsui-Jung
 ; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
 ; FILE REFERENCE: 201573
 ; CURRENT APPLICATION NUMBER: US/09/344,882
 ; CURRENT FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: US 60/090,717
 ; PRIOR FILING DATE: 1998-06-26
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: PatentIn Ver. 2.2

SEQ ID NO 24
LENGTH: 534
TYPE: PRT
ORGANISM: Arabidopsis Thaliana
US-09-344-882-24

Query Match 23.4%; Score 559.5; DB 4; Length 534;
Best Local Similarity 32.7%; Pred. No. 1.1e-43;
Matches 160; Conservative 76; Mismatches 220; Indels 33; Gaps 12;

QY 6 QNVVNGEMKSSVNV--QIETLSPIDSSLSGFVPMRTREVDHAKAGREALP--ANALTV 61
DB 55 QLLIGRFVDAVSGKTFFPLDPNGEVIAQVSEGDADVNAVAAARKKAFDEGPMFKMA 114
QY 62 YERAOYLHKAADIIEERDKEIATVLAKEISKAYNASV--TEVVRADLIYYAAEEGRIST 120
DB 115 YERSKILRFADLIEGHNDIEALEETWMDNGKPEGSAQIEVPMIAVFVYVYGMADKI-- 172
QY 121 SADEGGKMDASTGKHLAVIRROPVGIIVLAIAPVYVNLGSGKIAIPALIGANVMEKPT 180
DB 173 ---HGWTMGDGPVHHVQTL--HEPIGVAQGIIPWNPFLMLSMWLGDPALACGNTVVLKTA 228
QY 181 QSSVSGVLAKAPAEAGLPAGVFNITTGKSGEIGDYIVHEEVNFIINFGSTPVGQRIK 240
DB 229 QPPLSALIVGKLHHEAGLPDGVNIVSGFGATAGAIAASHMDVKVAFGSTDVGKILIE 288
QY 241 LA---GMRPIMELGKDAAGIYLADADLNNAAQIVAGAYDVSQGRCTAIKRVLVEEVA 297
DB 289 LASKSKLKVATLELESHEHSFPCEDADVQAVEIAHPALFPNQGCCCCGSGRTFVHERY 348
QY 298 DELAEKISENVAKLSVGDPPDNATVT--PVIDNSADPIESLV---VDA---RQKAKELN 350
DB 349 DEFVEKAKARALKRNVGDPFSGIEGPGVDSEQFKILKYIKHGEVAGATLQAGGDRIG 408
QY 351 EKRQDGRLLTPGLFPHVITLDMKLAMEEPFGPIPIIRVQDAEAAVIAKSDPGLSSVP 410
DB 409 S---KGYVQPTVFSVDKMDMLATDEIFGPVQTLIKFKDDEVLARANNRSRGLAAYV 465
QY 411 TRDFQKAPDIANKLEVGTVHNNKTGRGPDP---NPPFLGKSGAGVQGIIRYSIETAMTN 466
DB 466 TQNLDTAHLMLRALRGTWVINC-----DVLDAIIPFGGYKKSIGREKGIYSLANNYQ 520
QY 467 VKSIVLDMK 475
DB 521 VKAVVTSLK 529

RESULT 73
US-09-538-092-1173
Sequence 1173, Application US/09538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Glot, Loic
APPLICANT: Manefield, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538,092
CURRENT FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CurataseqFormatter Version 0.9
SEQ ID NO 1173
LENGTH: 535
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: Polypeptide Accession Number P51649
US-09-538-092-1173

Query Match 23.2%; Score 555; DB 4; Length 535;
Best Local Similarity 29.7%; Pred. No. 2.8e-43;
Matches 141; Conservative 97; Mismatches 220; Indels 16; Gaps 7;

QY 7 NYVNGEMKSSVNVQIETLSPIDSSLSGFVPMRTREVDHAKAGREALPAAALTVYERAO 66
DB 63 SFVGGRWLPAAATFPVQDPAGGAALGVADCGVEEAAVAAYAPFRKREVSARKSS 122
QY 67 YLHKAADIIEERDKEIATVLAKEISKAYNASVTEVVRADLIYYAAEEGRISTSDSG 126
DB 123 LIRKNTNLMQKDLAKITLAESEKPLKEAHGELVLSAFLEWFSERARRV-----YGD 177
QY 127 KMDASTGKHLAVIRROPVGIIVLAIAPVYVNLGSGKIAIPALIGANVMPPTQGSVSG 186
DB 178 IIHFAKORRLALVKQPIGVAVAIVTPNMFPSAMITRKGAALLAAGCTIVVPAEDTPESA 237
QY 187 LVLAFAFAAGLPAGVFNIT--TGSGEIGDYIVHEEVNFIINFGSTPVGQRIKLA- 242
DB 238 LALAEIAQAGIPSGVNVVIPCRRNAKEVEALCTDPLVKSISTGSTTGKILHHA 297
QY 243 -GMRPIMELGKDAAGIYLADADLNNAAQIVAGAYDVSQGRCTAIKRVLVEEVADELA 301
DB 298 NSVKVSMELGGLADPFIQDSANVDQAVAGMASKFRNTGQTCVCSNQLVQRGTHDAFV 357
QY 302 EKISENVAK-LSVGDPPDNATVT--PVIDNSADPIESLVVVAROKAKELNEFKR---DG 356
DB 358 KAFNAAMKKNLRVNGFPEGTTQGLINERKAVEKEKVNDVAVSGAIVVVGKRGHOLG 417
QY 357 RLITPGLFDHVTLDMKLAMEEPFGPIPIIRVQDAEAAVIAKSDPGLSSVFTPRDQK 416
DB 418 NFEFPITLLCNVTQDMLCTHEETFGIAPVIFDTEEEBAIAIANADVLGAGFYQSDBAQ 477
QY 417 AFDINKLEVGTVHNNKTGRGPDPNPPFLGKSGAGVQGIIRYSIETAMTNYSI 470
DB 478 IWRVAEQLVEGVAGVNEGLISSV--CPFGVAKQSGLREGSKYGIDEVLEIKTV 530

RESULT 74
US-09-328-352-7905
Sequence 7905, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7905
LENGTH: 564
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7905

Query Match 23.2%; Score 555; DB 4; Length 564;
Best Local Similarity 30.8%; Pred. No. 3.1e-43;
Matches 151; Conservative 93; Mismatches 208; Indels 38; Gaps 13;

QY 9 VNGE---W---KSSVNOIEILSP---IDDSLSGFVPMRTREVDHAKAGREALP---P 54
DB 56 VDGELKLMAGATKTLSPITWIQPPGSGIQVYELSGYPVNGESESDEALAAVRAVANNRG 115
QY 55 AMAALTVERAOYLHKAADIIEERDKEIATVLAKEISKAYNASVTEVVRADLIYYAAE 114
DB 116 EWPMAKVERIAQCMQNFQRMVEGRDLIKLIMEIGKSLADSEKEPRTITTYMR-----Q 171
QY 115 GIRLSTSADEGKMDASTGKHLAVIRROPVGIIVLAIAPVYVNLGSGKIAIPALIGANV 174
DB 172 TIDALKDIDNANSRFVIAEGTIGQIRTRPLVGLCMGVNPLNTPATLLIPALMGNTI 231
QY 175 MFKPTQGSVSGVLAKAPAEAGLPAGVFNITTGKSGEIGDYIVHEEVNFIINFGSTPV 234

```
Db      232 IFRPQGTLLFPLLEAFRDS-FPKVINTIYAPGSLVPHLLASQINVLLAISSKV 290
Qy      235 GORIGKLAKMRP-----IMLEIGKDGAGIVLADADIDNAKOIVAGANDYSQORCTAK 288
Db      291 ADHLKK---QHPKSHRLRALGLDAKVAAILLPDADIDLTVKECLGALSFNQORCTALK 347
Qy      289 RVLVBEVADLEKXISENVAKLSVGDPPD-NATVTPVIDNSADFIESIYVARQKGA 347
Db      348 MLMVHRSIADIEFVNRLTTEBLAKLVGMPWEKGVSTPPLPGMHRATVTEVEDAVAKGAP 407
Qy      348 ELN-----EFKRDRLTPGLFDHYTLDMLKLAWEPEPPIITIRVCAEBAVAIANKSDP 403
Db      408 VVNPBEGEPCK--TMFPYPAVVPVTEGMRIRYREGEQGPVPVPAVYDDIETVLDYVTTSDH 465
Qy      404 GLQSSVFTRDFOKAFDIANKL--EYGVTHINNTGRGPDNPPLGLKSGAGVQGI RYSTI 461
Db      466 GQGVSIQSDPAQIGHLVDTLVHGVCRVINQCQGRGPDVPFPGGRKDSAGTLSVHDAL 525
Qy      462 EAMTNVKSIV 471
Db      526 RAF-SIRSMI 534
```

RESULT 75

```
US-09-351-150A-9
; Sequence 9, Application US/09351150A
; Patent No. 6538177
; GENERAL INFORMATION:
; APPLICANT: Duvick, J
; TITLE OF INVENTION: Compositions and Methods for Fumonisin Detoxification
; FILE REFERENCE: 035718/167676, 5718-62
; CURRENT APPLICATION NUMBER: US/09/351,150A
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Bacterium 2412.1
US-09-351-150A-9
```

Query Match 23.2%; Score 554.5; DB 4; Length 468;

Best Local Similarity 30.7%; Pred. No. 2,5e-43; Matches 150; Conservative 91; Mismatches 205; Indels 43; Gaps 13;

```
Qy      4 EYQNVVNGEKKSSVNOIEILSPIDSSIGF-VPMATREEVDMHMKGREALPMAALTVY 62
Db      3 EYKLLIDRLVAGATTMSVINPATETPLVIDCPRADRDQDEAVAAERAFQSMRATLTLE 62
Qy      63 ERAQYIHKADIIERDEKEIATVLAKEISKAYNAVTEVVRTADLIRYAEIGIRLSTSA 122
Db      63 QRATINAIADAEADQSALARLTQEGKPLADAMGEIYASAAFRY-----FTSL 114
Qy      123 DEGGKM--DASTGHKLAVIRQPVGYVLAIPYNYPNLSGSKIAPALLIGANNVMKPPPT 180
Db      115 DLPPRVVRDPAATG--VEVHRRPLGVGCVIPNNFPMILMAFKIPALLAGNTVILKPPA 172
Qy      181 QGSVSGVLAKAFABAGLPAGVFNITTGSGEISGDYVHEHEVNFNFSTGTPVGORI-- 238
Db      173 TTPLTALRFG-ALVKDITVPGVNITITD--ADDLGAENTRHPGIRKISFTGSTGTGKKVMA 230
Qy      239 GKLAGMRPIMLEIGKDGAGIVLADADIDNAKOIVAGANDYSQORCTAIRVLVEEVAD 298
Db      231 GAAGELKRISLEIGNDALIVLDVDPEKEVAPRVFAAMQNAQOVCTIAAKRIYVHESLYE 290
Qy      299 ELAEKISENVAKLSVGD-----PFDN----ATVPYIDNSADFIESIYVDARQK 345
Db      291 AMCEEPQQLARIVVGGLEQGVOMGPLNRRQPEKVLGLIERARD--GRITAGGRKG 348
Qy      346 AKELNEFKRDGRLLTPGLFDHYTLDMLKLAWEPEPPIITIRVDAEBAVAIANKSDPGL 405
Db      349 DK-----GYPIEPIIVDIAEGQVLVDEEQGPVMPVIRYSDPVDAVRANASPYGL 400
```

```
Qy      406 QSSVFTRDFOKAFDIANKLEVGTVHINNTGRGPDNPPLGLKSGAGVQGI RYSTI EAMT 465
Db      401 GGSIMSRNVVRMSIADAEAGSVWVVKHADVQPD-LPLGAKFSQMSGLBEGGLHFT 459
Qy      466 NVKSIYLDM 474
Db      460 QVQ--VLNM 466
```

Search completed: August 28, 2005, 10:29:22
Job time : 50 secs

This Page Blank (uspto)